JGR Oceans

RESEARCH ARTICLE
10.1029/2019JC015128

Ecoregions in the Mediterranean Sea Through the Reanalysis of Phytoplankton Functional Types and Carbon Fluxes

S. Ciavatta[1,2], S. Kay[1], R. J. W. Brewin[1,3], R. Cox[4], A. Di Cicco[5], F. Nencioli[1,2], L. Polimene[1], M. Sammartino[5], R. Santoleri[5], J. Skákala[1,2], and M. Tsapakis[6]

[1]Plymouth Marine Laboratory, Plymouth, UK, [2]National Centre for Earth Observation, Plymouth Marine Laboratory, Plymouth, UK, [3]College of Life and Environmental Sciences, University of Exeter, Cornwall, UK, [4]AquaBioTech Group, Mosta, Malta, [5]CNR—Istituto di Scienze Marine, Rome, Italy, [6]Institute of Oceanography, HCMR, Anavyssos, Greece

Key Points:
- Ocean color phytoplankton functional types (PFTs) were assimilated into an ecosystem model of the Mediterranean Sea (1998-2014)
- The Mediterranean Sea can be subdivided in three ecoregions, based on the reanalyzed PFTs ratios to total chlorophyll
- The trophic and carbon sedimentation efficiencies are highest in the microphytoplankton-dominated ecoregion

Supporting Information:
- Supporting Information S1

Correspondence to:
S. Ciavatta,
s.ciavatta@pml.ac.uk

Citation:
Ciavatta, S., Kay, S., Brewin, R. J. W., Cox, R., Di Cicco, A., Nencioli, F., et al. (2019). Ecoregions in the Mediterranean Sea through the reanalysis of phytoplankton functional types and carbon fluxes. *Journal of Geophysical Research: Oceans*, *124*. https://doi.org/10.1029/2019JC015128

Received 6 MAR 2019
Accepted 31 AUG 2019
Accepted article online 6 SEP 2019

Abstract In this work we produced a long-term reanalysis of the phytoplankton community structure in the Mediterranean Sea and used it to define ecoregions. These were based on the spatial variability of the phytoplankton type fractions and their influence on selected carbon fluxes. A regional ocean color product of four phytoplankton functional types (PFTs; diatoms, dinoflagellates, nanophytoplankton, and picophytoplankton) was assimilated into a coupled physical-biogeochemical model of the Mediterranean Sea (Proudman Oceanographic Laboratory Coastal Ocean Modelling System-European Regional Seas Ecosystem Model, POLCOMS–ERSEM) by using a 100-member ensemble Kalman filter, in a reanalysis simulation for years 1998–2014. The reanalysis outperformed the reference simulation in representing the assimilated ocean color PFT fractions to total chlorophyll, although the skill for the ocean color PFT concentrations was not improved significantly. The reanalysis did not impact noticeably the reference simulation of not assimilated in situ observations, with the exception of a slight bias reduction for the situ PFT concentrations, and a deterioration of the phosphate simulation. We found that the Mediterranean Sea can be subdivided in three PFT-based ecoregions, derived from the spatial variability of the PFT fraction dominance or relevance. Picophytoplankton dominates the largest part of open ocean waters; microphytoplankton dominates in a few, highly productive coastal spots near large-river mouths; nanophytoplankton is relevant in intermediate-productive coastal and Atlantic-influenced waters. The trophic and carbon sedimentation efficiencies are highest in the microphytoplankton ecoregion and lowest in the picophytoplankton and nanophytoplankton ecoregions. The reanalysis and regionalization offer new perspectives on the variability of the structure and functioning of the phytoplankton community and related biogeochemical fluxes, with foreseeable applications in Blue Growth of the Mediterranean Sea.

1. Introduction

The composition of the phytoplankton community has a profound influence on the functioning of marine ecosystems, impacting ecosystem health (e.g., eutrophication and harmful algae blooms, e.g., Hallegraeff, 2003), services (e.g., aquaculture and fishery, e.g., Cury et al., 2008), and climate feedbacks through carbon sequestration (e.g., Bopp et al., 2005). Phytoplankton impact marine biogeochemical fluxes and trophic networks, in relation to their size (e.g., picophytoplankton (PIC), nanophytoplankton (NNP), and microphytoplankton (MIC) have different affinities for nutrients and only the latter are palatable to fishes) and to their functional characteristics (e.g., diatoms, typically large in size, can assimilate silicate, while nanosized coccolitophores impact the cycling of calcium; Chisholm, 1992). Phytoplankton functional types (PFTs) are defined as groups of phytoplankton that have a distinct function in the ecosystem (Baretta et al., 1995).

PFTs can be investigated by means of ocean color observations (IOCCG, 2014). A variety of algorithms have been proposed to derive PFTs biomass and production from satellite observations (Kostadinov et al., 2017). However, these data sets are limited to the first optical depth of the ocean, and the derivation of vertical structure and of fluxes other than primary production is highly uncertain (Lee et al., 2015). State-of-the-art ecosystem models can describe PFTs and the related biogeochemical and trophic fluxes, but the skill of model simulations is often poor, mainly due to the difficulty of formulating and parameterizing PFT processes (Shimoda & Arhonditsis, 2016). Merging PFT ocean color and ecosystem models through data assimilation

was proposed as a novel approach to estimate better the phytoplankton community structure and related fluxes in marine ecosystems (Ciavatta et al., 2018). Better simulations of diatoms, dinoflagellates, NNP, and PIC biomass, as well as estimates of related biogeochemical fluxes, were delivered in both a reanalysis simulation (Ciavatta et al., 2018) and (pre)operational predictions (Skákala et al., 2018) in the North West European Shelf Sea ecosystem, through the assimilation of a regional ocean color PFT product (Brewin et al., 2017).

In the Mediterranean Sea, regional ocean color algorithms and products were used to characterize the spatial-temporal variability of PFTs (Di Cicco et al., 2017; Navarro et al., 2014, 2017; Sammartino et al., 2015; Uitz et al., 2012; Volpe et al., 2007). Observation-based analysis showed that small-sized groups, either PIC (Magazzù & Decembrini, 1995) or NNP (Uitz et al., 2012), can dominate the primary production in the Mediterranean Sea, although MIC production can also be relevant in some spring-blooming Mediterranean areas (Uitz et al., 2012). Modeling studies have investigated the variability of primary production (e.g., Di Biagio et al., 2019; Lazzari et al., 2012, 2016), but they only focused on the phytoplankton community as a whole.

Model and ocean color total chlorophyll were used to define ecoregions in the Mediterranean Sea (Ayata et al., 2018; Basterretxea et al., 2018; d'Ortenzio & Ribera d'Alcalà, 2009; Mayot et al., 2016). Ecoregions can be defined as areas characterized by distinctive and rather homogeneous ecosystem features, such as the phenology of the total phytoplankton community, as conducted in the above cited papers. The above total chlorophyll-based regionalizations, along with regionalizations based on hydrodynamics and biogeochemistry, have been recently integrated as layers of a "consensus" map, aiming to support the understanding of the ecosystem functioning as well as the implementation of the Marine Strategy Framework Directive in the Mediterranean Sea (Ayata et al., 2018). To the authors' knowledge, a regionalization based on PFTs from either ocean color or models has not been carried out yet, despite PFTs having profound effects on the ecosystem features, such as biogeochemical fluxes and trophic webs.

The objective of this work is to define PFT-based ecoregions in the Mediterranean Sea and to characterize their carbon fluxes, by merging information on PFTs from model simulations and ocean color observations.

To achieve this objective, we assimilated a regional ocean color product for PFTs (Di Cicco et al., 2017) into an assimilative modeling system for the Mediterranean Sea. This system uses the coupled physical-biogeochemical model Proudman Oceanographic Laboratory Coastal Ocean Modelling System (POLCOMS)-European Regional Seas Ecosystem Model (ERSEM) of the Mediterranean Sea developed by Holt et al. (2008) and updated by Kay and Butenschön (2018), and we applied the PFT assimilation framework developed by Ciavatta et al. (2018). Previous applications of ERSEM to investigate the Mediterranean Sea biogeochemistry, in both assimilative and not assimilative simulations, include the works by Vichi et al. (1998), Allen et al. (1998, 2002, 2003), Petihakis et al. (1999, 2002, 2007), Triantafyllou et al. (2000, 2003), Zavatarelli et al. (2000), Blackford (2002), Pinardi et al. (2003), Kalaroni et al. (2016), and Tsiaras et al. (2017). A skill assessment was performed to evaluate the reanalysis adequacy in representing the average spatial variability of the phytoplankton community structure. The ecoregions have been defined by computing multiannual average of reanalyzed surface PFT fractions, assessing spatial patterns of the PFT dominance or relevance, and evaluating the spatial variability of vertically integrated carbon fluxes.

The paper is structured as follows. Section 2 describes the ecosystem model, the setup of the assimilation algorithm, the data, the metrics applied for skill evaluation, and the method applied to identify the different PFT ecoregions and to characterize selected carbon and trophic fluxes. The results are presented and discussed in section 3, where the skill of the reanalysis is evaluated by using (i) the assimilated ocean color chlorophyll products (section 3.1) and (ii) not assimilated in situ data set (section 3.2). Section 3.3 presents the PFT-based ecoregions, and section 3.4 compares their carbon and trophic fluxes. Concluding remarks and future applications are pointed out in section 4.

2. Material and Methods

2.1. The Ecosystem Model of the Mediterranean Sea

The ecosystem dynamics of the Mediterranean Sea (Figure 1) are described by a three-dimensional physical-biogeochemical model. This consists of three online coupled submodels (see Figure 2). The POLCOMS (Holt

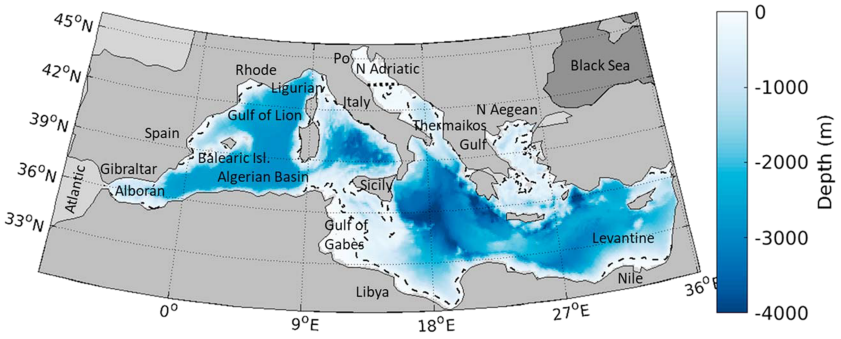

Figure 1. Domain and bathymetry of the Mediterranean Sea model. The western boundary of the model extends into the Atlantic as a buffer area, but results will be only shown for the Mediterranean area in the map, that is, from 5°W. The dashed line represents the 200-m isobaths; the dotted line represents latitude 43.5°N, here delimiting the North Adriatic.

et al., 2008; Holt & James, 2001) describes the hydrodynamics. This provides the physical forcing to the pelagic biogeochemical submodel, namely, the ERSEM (Baretta et al., 1995; Butenschön et al., 2016). The third submodel is the ERSEM benthic biogeochemical model (Blackford, 1997; Butenschön et al., 2016). The two ERSEM submodels are coupled at the same temporal and spatial resolution as the physical model to capture the effects of the three-dimensional hydrodynamics on the biogeochemical cycles (Holt et al., 2004). The model grid has a horizontal resolution of 0.1° (approximately ~11 km in the meridional direction and between 9.3 and 7.8 km in the zonal direction at the latitude of the study region) and 42 sigma coordinate levels in the vertical.

2.1.1. The Physical Submodel: POLCOMS

The physical model POLCOMS (Holt & James, 2001) is a three-dimensional, finite difference, primitive equation model formulated in spherical-polar coordinates on an Arakawa B-grid. The model includes an advection scheme with stability and conservation properties (James, 1996), a vertical turbulence model (General Ocean Turbulence Model; Burchard et al., 1999), and calculation of horizontal pressure gradients. POLCOMS was configured for the Mediterranean Sea by Holt et al. (2008).

2.1.2. The Pelagic Biogeochemical Submodel: ERSEM

The biogeochemical processes are described by the ERSEM (Baretta et al., 1995), version 15.06 (Butenschön et al., 2016), as in the PFT assimilative model by Ciavatta et al. (2018). ERSEM uses a functional type approach to model the dynamics of the low trophic levels of the ecosystem. Primary producers are split into four PFTs, including two size-based types (PIC and NNP), dinoflagellates, and large diatoms (which, summed up, represent the model MIC). Each of these PFTs is represented in terms of its content of chlorophyll, carbon, nitrogen, phosphate, and (for diatoms only) silicate. Three functional types of zooplankton (mesozooplankton, microzooplankton, and heterotrophic nanoflagellates) prey on the phytoplankton, bacteria, and particulate organic matter as a function of their size. A bacterial functional type drives the microbial loop, the production and recycling of dissolved organic matter in labile and recalcitrant forms, and the regeneration of inorganic nutrients in the water column (Hansell, 2013; Polimene et al., 2006). In the functional types, the stoichiometric ratios of nutrients-to-carbon and chlorophyll-to-carbon (in the PFTs) vary dynamically (Baretta-Bekker et al., 1997; Geider et al., 1997). The model simulates the dynamics of five inorganic dissolved nutrients (carbon, nitrate, ammonium, phosphate, and silicate) and dissolved oxygen. Nitrogen undergoes nitrification in the water column, as a function of available ammonium, temperature, oxygen, and pH (Blackford & Gilbert, 2007; Butenschön et al., 2016). Denitrification is modeled in the sediments only (Butenschön et al., 2016).

ERSEM coupled to POLCOMS was used in the Mediterranean Sea by Kay and Butenschön (2018) and Kay et al. (2018), by using the ERSEM parameterization of Butenschön et al. (2016). However, here the maximal PFT carbon-to-chlorophyll ratios were set equal to the values in the comparable Mediterranean model by Lazzari et al. (2012, 2016) and Teruzzi et al. (2014, 2018); i.e., 0.02 mg Chl/mg C). In fact, the original ERSEM values were calibrated for the global ocean and led to ocean color overestimation in the Mediterranean Sea in our preliminary sensitivity simulations (not shown).

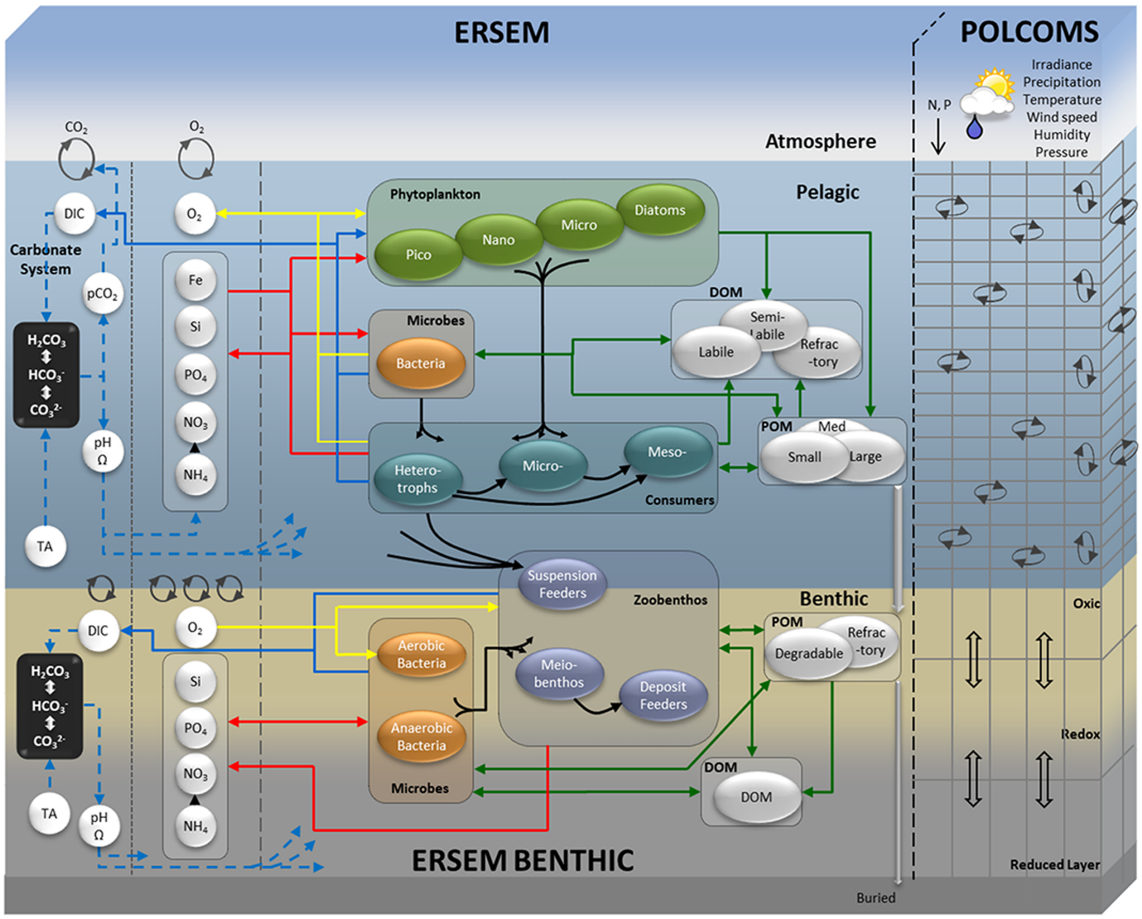

Figure 2. Schematic of the coupled physical-biogeochemical submodels of the Mediterranean Sea, including the list of the forcing functions. ERSEM = European Regional Seas Ecosystem Model; POLCOMS = Proudman Oceanographic Laboratory Coastal Ocean Modelling System.

2.1.3. The ERSEM Benthic Submodel

The benthic submodel is the ERSEM benthic model (Blackford, 1997), as described in Butenschön et al. (2016). In the configuration applied here, the submodel includes 35 biogeochemical variables, subdivided into five living functional types (including three zoobenthos types, and aerobic and anaerobic bacteria), along with particulate matter and dissolved organic and inorganic nutrients. The fluxes at the sediment-water interface are determined by sedimentation of organic matter and diffusion of inorganic nutrient across the seabed.

2.1.4. Atmospheric Forcing, Boundary Constrains, and Initial Conditions

Atmospheric forcing was provided by the ERA-Interim data set (Dee et al., 2011), using 6-hourly values of surface winds, sea level pressure, temperature and relative humidity, and daily accumulation of precipitation and shortwave and longwave radiation fluxes. The resolution of the surface data was 0.75°. Ocean boundary conditions at the western boundary were taken from a run of the same model configured for a wider domain (Kay et al., 2018) using the same surface forcing: This provided daily values of temperature, salinity, currents, nitrate, phosphate, silicate, dissolved inorganic carbon (DIC), and bioalkalinity. Climatological values for river inputs of fresh water and nutrients were taken from GlobalNEWS2 (Mayorga et al., 2010). River nutrient concentrations were set constant; riverine freshwater discharges varied seasonally using historic data from Dai et al. (2009) to set the variation. Climatological inputs for the North Adriatic (NAd) rivers, including the Po River, were derived from data in Cozzi and Giani (2011). Atmospheric depositions of bioavailable nitrogen and phosphorus were represented as in Lazzari et al. (2016), that is, temporally constant and spatially different in the eastern and western basins, with values derived from Ribera d'Alcalà et al. (2003).

The initial conditions for nitrogen were derived to match the average subbasin biogeochemical-ARGO profiles in de Fommervault et al. (2015). Phosphate initial conditions were set equal to 1/25 of the

nitrogen values, that is, setting the initial *N*:*P* ratio equal to an average Mediterranean value of 25 (comparable, e.g., to Pujo-Pay et al., 2011). The initial conditions for the other biogeochemical and physical variables were derived from a model simulation on a wider model domain (Kay et al., 2018). A spin-up of 20 years was performed before starting the reference and assimilative simulations, having verified that this was sufficient to stabilize the patterns of the nutrients profiles.

2.2. The Assimilation System

The assimilation system is described in full in Ciavatta et al. (2016, 2018). It uses the ensemble Kalman filter (EnKF; Evensen, 1994). This is a sequential assimilation method, which starts by randomly sampling an ensemble of N state vectors $x_0^{a(l)}$ ($l = 1, 2, \ldots, N$) from an initial probability density function for the model variables. Each ensemble member, that is, state vector, is propagated in time using the nonlinear model equations during the "forecast step" that provides the EnKF "forecasts" $x_i^{f(l)}$, that is, the reanalysis states just before the subsequent assimilation step.

At time i, the forecast state x_i^f and the forecast uncertainty P_i^f are defined from the mean value and the covariance matrix of the N forecasted members. When at time i a vector y'_i of observations of the model output $y_i = H\left[x_i^{f(l)}\right]$ becomes available, the assimilation scheme updates (i.e., "corrects") the forecasted states $x_i^{f(l)}$, in the EnKF "analysis" step. This step scales the forecast-to-data mismatches, by balancing the uncertainty in the model (P_i^f) and in the observations (R_i), and it provides the analyzed ensemble $x_i^{a(l)}$. The difference $x_i^{a(l)} - x_i^{f(l)}$ is the "increment" (i.e., the correction) of the state vector in the analysis step. The ensemble $x_i^{a(l)}$ is the initial condition used to simulate a new ensemble forecast for time $i + 1$, in a sequential procedure that estimates the evolution of the model variables over the time window spanned by the assimilated data.

Our assimilation system uses the Evensen (2003) version of the EnKF, which includes localization of the analysis and perturbation of the assimilated observations (see also Hu et al., 2012; Natvik & Evensen, 2003; Storto et al., 2013). The observational operator $H\left[x_i^{f(l)}\right]$ links the four state variables representing the ERSEM phytoplankton types to the respective ocean color products (Ciavatta et al., 2018). Observations and model states are log-transformed prior to the analysis, to guarantee positivity of the solutions (Janjić et al., 2014), as in the applications by Nerger and Gregg (2007), Ciavatta et al. (2016, 2018), Ford and Barciela (2017), and Skákala et al. (2018).

The PFT assimilation scheme was set up as in Ciavatta et al. (2018). We used the EnKF with an ensemble size of $N = 100$ members. We included in the analyzed state vector 35 out of the 51 ERSEM pelagic state variables, to keep the analysis computationally affordable, in analogy with Ciavatta et al. (2018). Coherently with the paper's objective to reanalyze the plankton community structure, we included in the state vector all the 27 variables related to the phytoplankton, zooplankton, and bacteria functional types, as well as eight variables representing DIC, labile, and refractory dissolved organic matter and small particulate matter. The remaining 16 variables were updated through the model equation during the simulation runtime ("forecast" step; e.g., Pradhan et al., 2019). The radius of the localized analysis was set to be spatially variable as a function of the bathymetry. In particular, we increased the "resolution" of the analysis from oceanic toward coastal waters, by setting a radius of 100 km for grid points where the bathymetry is deeper than 2,000 m, 50 km for bathymetry between 50 and 2,000 m, and 25 km for bathymetry shallower than 50 m (Ciavatta et al., 2016).

The model error is accounted for by random perturbations of the model forcing, namely, the surface solar irradiance, thus inducing fluctuations in the underwater light field that drives photosynthesis (Ciavatta et al., 2016). A Gaussian perturbation with standard deviation equal to 20% of the irradiance value is added during the model forecast step. Furthermore, at the first assimilation step of each year, the model error is added to all the variables undergoing the analysis, as white noise drawn from a distribution of pseudo-random fields with error equal to 10% of the value of the variables. The error is lowered to 1% for those variables that have relatively high mean values (DIC, small particulate matter), to avoid divergences in the concentrations of the largest pool in the model (Ciavatta et al., 2011).

The ensemble was initialized by perturbing the output of the spin-up reference model simulation for September 1997, that is, the first month for which ocean color products were available for

assimilation. The perturbation applied Gaussian pseudo-random fields with error equal to 30% of the value of the variables.

The observational error was derived from values of the bias and root-mean-square deviation (RMSD) of the PFT ocean color product (see section 2.3). The variance of the unbiased PFT concentrations regridded onto the surface model domain and was computed following the procedure described in the Appendix of Ciavatta et al. (2016).The PFT assimilation approach was developed in the framework of the EU Copernicus Marine Environment Monitoring Service (CMEMS) - Service Evolution Project TOSCA, and was applied with the operational CMEMS model of the North West European Shelf-Sea by Skákala et al. (2018).

The reanalysis simulation was run on the U.K. Natural Environment Research Council (NERC) High Performance Computing facility "ARCHER," using 7200 CPUs and ~10 mega allocation units and producing ~10 Tb of output.

2.3. Data

The assimilated ocean color PFT chlorophyll products were developed by Di Cicco et al. (2017) using regional PFT algorithms for the Mediterranean Sea. These algorithms are empirical relationships developed following a global approach (Hirata et al., 2011) based on the relation between phytoplankton abundance and the trophic status of the environment (Brewin et al., 2010, 2011; Chisholm, 1992; e.g., Devred et al., 2006; Uitz et al., 2006). These algorithms resulted from the existing covariability between in situ PFT groups (expressed as chlorophyll fraction) and the corresponding $\log_{10}$-transformed in situ total chlorophyll *a* concentration, taking account of the lognormal distribution of this pigment (Chisholm, 1992; Hirata et al., 2011). Following previous works carried out for the global ocean (Barlow et al., 1993; Gieskes et al., 1988; Uitz et al., 2006), the contribution of each PFT to the Mediterranean assemblage in terms of chlorophyll concentration was computed by performing a multiple regression analysis between in situ total chlorophyll *a* and seven diagnostic pigments, markers for the main algal groups (Vidussi et al., 2001). The final in situ PFT estimation formulas and the regional satellite algorithms, with their references, are fully described in Di Cicco et al. (2017).

The regionally tuned satellite PFT algorithms were applied to a Mediterranean multisensor (Medium-Resolution Imaging Spectrometer(MERIS), Moderate Resolution Imaging Spectroradiometer-Aqua (MODIS), and Sea-viewing Wide Field-of-view Sensor(SeaWIFS) chlorophyll *a* satellite time series (from September 1997 to August 2015) available from Copernicus Marine Environment Monitoring Service (CMEMS, OCEANCOLOUR_MED_CHL_L3_REP_OBSERVATIONS_009_073 product). These chlorophyll data are produced by the CMEMS OC TAC (Ocean Color Thematic Assembling Centre) using an ad hoc configuration of the ESA OC-CCI (European Space Agency-Ocean Color Climate Change Initiative, www.esa-oceancolour-cci.org) processor at high resolution in the Italian National Research Council Mediterranean processing chain (Volpe et al., 2012, 2019). The chlorophyll product used in Di Cicco et al. (2017) is a merged Case I–Case II product, in which the different optical properties of inshore and offshore waters are taken into account by applying two different regional algorithms: (i) the MedOC4 (Volpe et al., 2007) specialized for open ocean waters and (ii) the AD4 (D'Alimonte & Zibordi, 2003) for coastal complex water. The method of D'Alimonte et al. (2003) was used for the exact identification of the two different water types. The time series for this work includes daily values of chlorophyll *a* concentration quantified for four phytoplankton groups (PIC, NNP, diatoms, and dinoflagellates) and their associated errors (bias and RMSDs) at 2-km resolution. The ocean color PFT errors were estimated through model validation using an independent PFT validation in situ data set (Di Cicco et al., 2017) derived from the SeaBASS data set (Werdell & Bailey, 2005).

Using the same methods applied in Ciavatta et al. (2018), we computed unbiased concentrations and associated standard errors of the PFT concentrations. These were then projected onto the model grid. Finally, we computed 5-day composites centered on the last day of each month of years 1997–2015. These composites were assimilated on a monthly time step, and the results of the reanalysis are presented in this work for years 1998–2014.

To corroborate the simulated sea surface temperature (SST), we used data of the Group for High Resolution Sea Surface Temperature (GHRSST) global Level 4 AVHRR_OI product. This product uses optimal interpolation (OI) by interpolating and extrapolating SST observations from different sources, which include

satellite (Advanced Very High Resolution Radiometer [AVHRR]) and in situ platforms, i.e., ships and buoys (Reynolds et al., 2007). The original GHRSST global Level 4 AVHRR_OI data at daily temporal resolution and 0.25° spatial resolution were averaged to monthly values and interpolated to the model grid for comparison with the model output. The in situ biogeochemical observations used to evaluate the reanalysis skill were obtained from publicly available data sets. Temperature, salinity, nitrate, phosphate, and chlorophyll data were provided through OGS/NODC infrastructure (http://nodc.ogs.trieste.it); we used data of biogeochemical-ARGO profiles for nitrate and oxygen (http://doi.org/10.17882/42182#58265) as well as for salinity, temperature, and chlorophyll (BOPAD-prof, Barbieux et al., 2017).

In situ observations of PFTs were obtained from Di Cicco et al. (2017), in turn, derived from the SeaBASS data set (Werdell & Bailey, 2005).

2.4. Skill Assessment and Similarity Evaluation

The skill of the assimilation of the ocean color PFTs (reanalysis for 1998–2014) was assessed in comparison with a reference simulation without assimilation and in comparison with assimilated ocean color products and independent in situ observations, by using quantitative statistical metrics, as in Ciavatta et al. (2018). The performance of the reference and reanalysis output (y) in matching the assimilated composites of ocean color PFT concentrations (y'_c) was evaluated by computing spatial correlation and RMSD at each month of the simulation years and then averaging in time. The skill of the reanalysis was evaluated for the output of both the forecast and analysis steps of the assimilation algorithm (see section 2.2 for the definition of these steps). In particular, the assessment of the analysis skill used the same ocean color products that were assimilated monthly, while the assessment of the forecast used ocean color data that had not been assimilated yet. Although this is a common procedure in the evaluation of assimilative systems (see, e.g., Gregg et al., 2009), we stress that the not-assimilated ocean color data used for skill assessment cannot be considered independent from the assimilated ones, because of temporal correlations among subsequent ocean color fields retrieved from satellites. Similarly, also the nonassimilated in situ PFT observations cannot be considered fully independent because they were used to calibrate the assimilated ocean color PFT products (Di Cicco et al., 2017). The remaining in situ physical and biogeochemical observations are independent from the assimilated products.

The similarity between the reanalysis and ocean color average phytoplankton community was synthetized in a map scaled from 0 to 1. This map was derived by (1) computing and summing the absolute differences between the reanalysis and ocean color maps of the four PFT average fractions; (2) scaling the sums between 0 and 1, by dividing by the maximum sum; (3) computing the differences in (1) above minus the scaled sums. In the map, the value 1 represents the maximum similarity, while 0, the lowest.

Quantitative metrics were used to evaluate the skill of the reanalysis in matching independent in situ observations. Daily values of the variables in the reanalysis data set (y) were matched up point to point in space and time with the data (o). In particular, we computed and then showed in robust skill diagrams (Butenschön et al., 2016; Ciavatta et al., 2018) (a) the bias calculated as the median value of the reanalysis-to-observation mismatch, bias* = median($y - o$), normalized by the interquartile range of the data (IQR_o); (b) the unbiased median absolute error, MAE′ = median{abs[$y - o$ − bias*]}, normalized by IQR_o, and taken with the algebraic sign of the differences between the interquartile range of the output and the data, sign ($IQR - IQR_o$); and (c) the correlation coefficient. In these diagrams, the closer the point is to the center, the lower the error of the simulation (i.e., the median bias and the unbiased absolute error).

The coherence between in situ observations and model output was evaluated also by comparing their spatial distributions in bar plots, representing median and interquartile range. For discussion purposes, the matchups were subdivided in broad zones, encompassing coastal waters (bathymetry less than 200 m), open surface ocean and open deep ocean (with bathymetry more than 200 m and surface/deep thresholds of 10 m for chlorophyll and 200 m for nutrients). The in situ chlorophyll observations have a detection limit of 10^{-3} mg/L, and the same threshold was set to the matchup reanalysis data used in the skill assessment, for the sake of consistency.

2.5. Definition of the PFT Ecoregions and Carbon Fluxes

Three PFT ecoregions were defined by assessing the spatial variability of the reanalyzed MIC, NNP, and PIC fractions computed at the surface of the model domain, averaged in the years 1998–2014.

The PFT fractions were computed as ratios of each PFT chlorophyll to total chlorophyll. Diatoms and dinoflagellates fractions were summed up to define a region of dominance of MIC as a whole. The NNP and PIC regions were defined by weighting the ratios of these two PFTs (see Table 3 for details).

In order to characterize the different regions of the Mediterranean Sea, we have selected three simulated ecosystem fluxes that can be of interest for applications such as aquaculture or fisheries: (1) the ratio of gross primary production to community respiration (GPP:CR), which defines an autotrophic system when greater than 1; (2) the trophic efficiency, here approximated by the ratio between mesozooplankton and phytoplankton biomass (mesozooplankton is the highest trophic level represented in the model), which indicates how much of the primary production is transferred to the upper trophic levels (see, e.g., Kwiatkowski et al., 2019); and (3) the sedimentation efficiency, here defined as the flux of particulate organic carbon (POC) reaching the sediments, normalized by the GPP. All the fluxes, except the sedimentation efficiency, were computed by integrating the model output from the surface to the maximum depth of 200 m, which on average incorporates the entire euphotic zone, encapsulating the vertical distribution of phytoplankton in the Mediterranean Sea (Lavigne et al., 2015). Sedimentation efficiency was computed at the bottom of the water column and provides an indication of the POC available to the benthic filter feeders, for example, commercially relevant shellfish.

3. Results and Discussion

3.1. Skill in Matching the Ocean Color Products

The reanalysis reproduced the distributions of the PFT chlorophyll concentrations as inferred from ocean color observations in the surface waters of the Mediterranean Sea (Figure 3). On average, PIC and NNP reanalysis had the highest chlorophyll concentrations and displayed east-west differences, though the gradients were more evident in the satellite products. The MIC group (i.e., both diatoms and dinoflagellates) had lower average concentrations but high peak values and steep gradients in coastal areas (e.g., in the Adriatic Sea and Gulf of Gabès). The model overestimated ocean color PIC, leading to an overall overestimation of total chlorophyll. This holds in particular for the eastern basin, where NNP and diatoms were overestimated as well.

Crucially, the simulated average distribution of the phytoplankton community structure matched well the one inferred from the ocean color products, as shown in the maps of Figures 4 and 5. Both the model and the ocean color products indicate the following order of dominance of PFTs on the total chlorophyll in the Mediterranean Sea (Figure 6): (1) PIC, (2) NNP, (3) diatoms, and (4) dinoflagellates.

The PFT community structure has evident gradients in both the model and data fields (see PFT fractions in Figure 4 and Figure S1 in the supporting information) that reflects the gradients of the simulated surface nutrients (see section 3.2). PIC ratio is higher in the more oligotrophic eastern Mediterranean than in the western Mediterranean. Diatoms (and dinoflagellates in the reanalysis) dominate in the nutrient-rich coastal waters of the Adriatic Sea, the Gulf of Gabès, and the area surrounding the Nile estuary. NNP relevance is intermediate among the other PFTs (Figure 6): It is lower than PIC fractions in both the west and East Mediterranean according to both the satellite and reanalysis products (see Figures 4 and S1). However, the east-to-west and coast-to-open sea gradients are more evident in the satellite product that in the model simulation of the NNP and PIC ratios.

Figure 5 shows that the reanalysis and ocean color representation of the plankton community structure tends to be more alike in the open ocean than in coastal areas, where the normalized similarity index might be lower than 75%. In coastal areas, often MIC (i.e., diatoms plus dinoflagellates) dominate in both the reanalysis and ocean color product but with rather higher fraction values in the ocean color product (e.g., in the NAd, Gabès gulf and Nile delta, Figure 4). Similarity is lower than 75% also in the North West Mediterranean, North Aegean and Ionian Sea, where the reanalysis NNP fractions are noticeably lower than in the ocean color products.

3.1.1. Comparison With the Reference Simulation

The assimilation of ocean color PFT improved the reference model simulation of the phytoplankton community structure (i.e., PFT fractions, Figure 7a). This holds for the diatoms and dinoflagellates fractions in particular, which halved their normalized bias. The bias of the PIC also decreased slightly. NNP skill did not

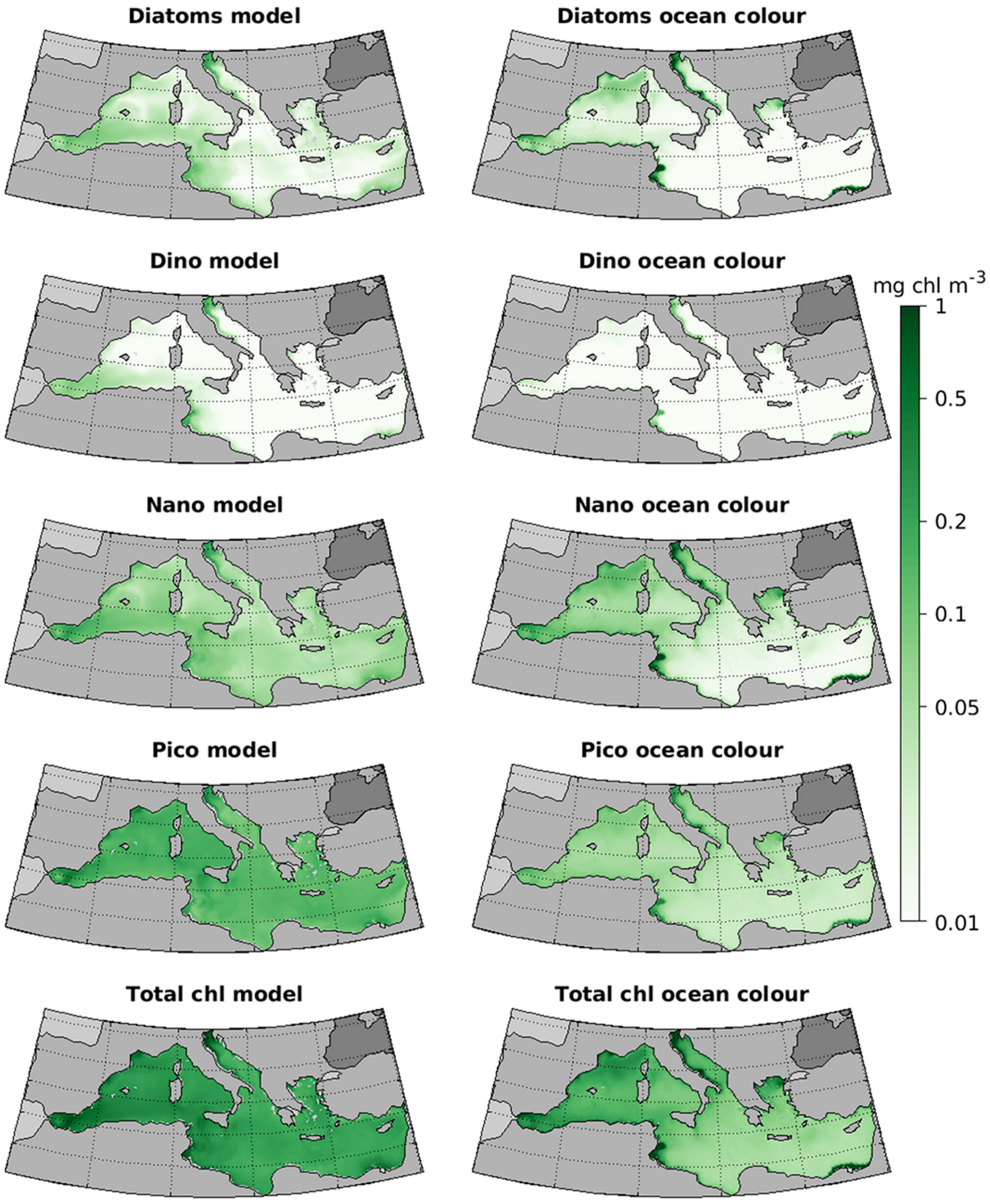

Figure 3. Comparison of the reanalysis versus ocean color estimates of phytoplankton functional types and total chlorophyll concentrations in the Mediterranean Sea. The maps show the surface average values in the period 1998–2014.

improve significantly, and their reference correlation remained negative, since the west-east gradient of the ocean color fractions remained unrepresented (Figure 4).

The reanalysis did not change significantly the reference skill for the PFT concentrations, excluding dinoflagellates that were slightly deteriorated (Figure 7b). This resilience of the skill is not related to a malfunctioning of the assimilative system but to feedbacks of the simulated ecosystem to the analysis improvements. In fact, the analyzes had better skill than the 1-month forecasts, as expected from a functional assimilative system, but the forecast skill remained comparable to the reference one (Table 1). The EnKF analysis increments shifted the model toward the ocean color gradients, by increasing strongly the biomass in the West Mediterranean (NNP) and along the coasts (dinoflagellates and diatoms) and decreasing the biomass in the east (NNP and PIC; see Figure S2 in the supporting information). However, these changes led also to proportional EnKF increments of nutrient content

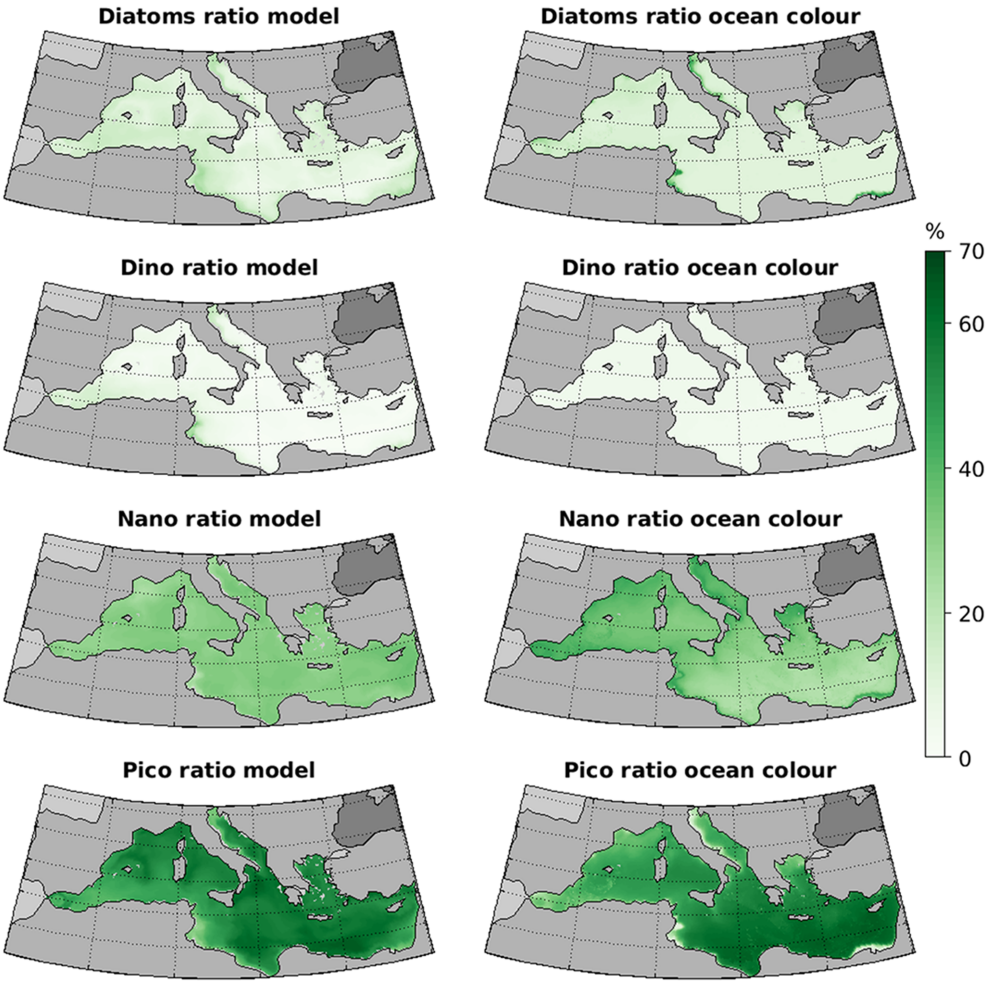

Figure 4. Comparison of the reanalysis versus ocean color estimates of the phytoplankton community structure in the Mediterranean Sea. The maps show the surface percentage fractions of the phytoplankton functional type chlorophyll with respect to total chlorophyll, in the period 1998–2014.

inside the phytoplankton (see phytoplankton phosphorus in Figure S2), which became a source of nutrients to the system. This source of nutrients exacerbated the model nutrient bias and sustained basin-wide growth of total phytoplankton, despite localized assimilative corrections of single PFTs (e.g., MIC increase in coastal areas). Therefore, the analysis was able to improve the reference spatial

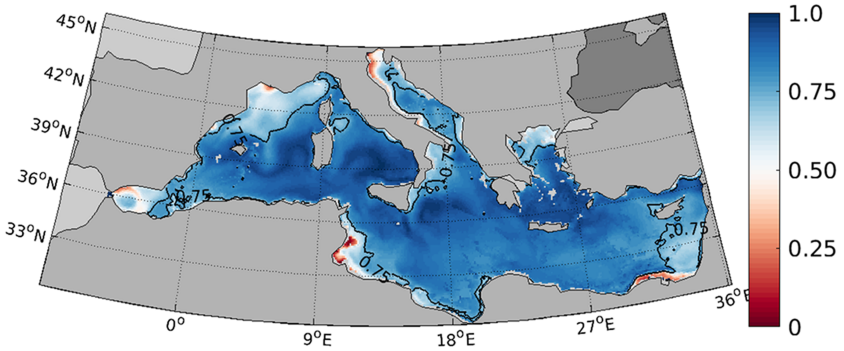

Figure 5. Map of similarity between the reanalysis and ocean color representations of the plankton community structure. The similarity was computed from the maps in Figure 4 as described in section 2.4. The value 1 indicates the maximum relative similarity, and 0, the lowest. The black line represents the 75% similarity level.

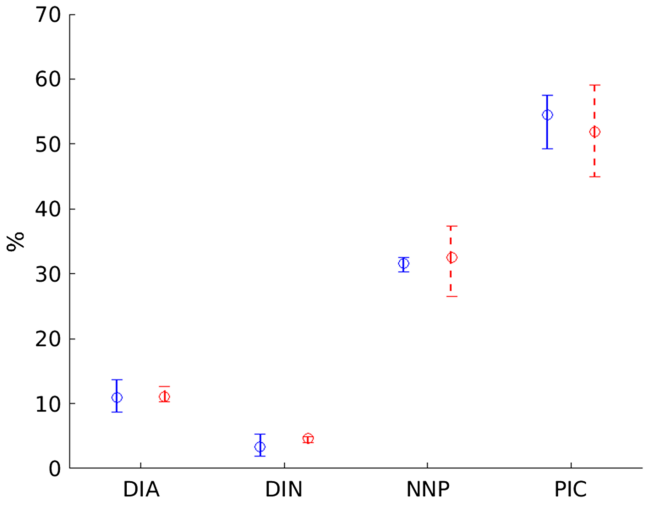

Figure 6. Spatial distributions of the average phytoplankton functional type fractions in the Mediterranean Sea from the reanalysis output (blue continuous bars) and ocean color (red dashed bars; surface, years 1998–2014) computed from Figure 4. The bars represent the first and third quartiles around the median values (circles). DIA = diatoms; DIN = dinoflagellates; NNP = nanophytoplankton; PIC = picophytoplankton.

distribution of the PFT fractions (Figure 7a) even though the absolute chlorophyll concentration of each PFT was not improved (Figure 7b).

3.2. Skill in Matching In Situ Observations

The reanalysis was relatively skilled in fitting the not-assimilated observations of physical and biogeochemical variables (Figure 8 and Table 2) collected at the sampling sites shown in Figure 9. Median bias and absolute errors were smaller than the variability of the observations for all the variables, except for oxygen, as indicated by the circles within the unit radius in Figure 8. The interquartile ranges of data and reanalysis overlapped for all the variables (Table 2) and the correlations for temperature, salinity, and chlorophyll were as high as 0.92, 0.66, and 0.44, respectively (Figure 8). Correlations for the other variables were relatively low. Nitrate reanalysis and data overlapped, though the simulations overestimated the observed distribution (Table 2). Similarly, phosphate observations were overestimated in the reanalysis: The positive bias and rightside position in Figure 8 reflects the higher median value and larger interquartile range in Table 2. The phosphate overestimation likely explains the overall model overestimation of chlorophyll concentrations (Figure 3). In fact, phosphate remains on average the limiting factor of primary production in our simulation, in agreement with previous experimental and modeling findings (e.g., Krom et al., 1991; Lazzari et al., 2016), implying overproduction where this nutrient was overestimated. However, we note that phosphate data were available for model assessment only in limited areas of the Mediterranean Sea.

The reanalysis captured the highest observed PFTs concentrations well (see the comparable values of IQ3 in Table 2, with the exception of the underestimated NNP) but the simulated distributions tended to be skewed toward low values (Table 2), leading to an overall negative median bias (Figure 8). We note that the number of in situ observations available for the assessment of PFTs was relatively low compared to the other variables.

The spatial patterns of the errors were different for the different variables, suggesting that the coupled physical-biogeochemical model did not have systematic regional failures (Figure 9). Errors of simulated temperature were higher in the area of the Gulf of Lion and Ligurian Sea (Figure 9). In the same area, also a comparison with the GHRSST SST confirms a larger model uncertainty (Figure S3 in the supporting information). This can be related to the model underestimation of the velocities of the Ligurian and Balearic currents and excessive summer stratification of the water column (Kay et al., 2018). The model simulated poorly other complex hydrodynamic features, such as the gyres in the North Tyrrhenian and South Adriatic and currents around the Sicily Island: See Figure 9 and Figure S3 in the supporting information. This uncertainty is probably a consequence of the coarse resolution of the model grid (0.1°) and of the meteorological forcing (0.75°) and has likely contributed to the bias of the model simulation of nutrients and PFT concentrations.

The error patterns of salinity were less systematic, since high and low error points coexisted in the same area; see, for example, the Algerian Basin and the Tyrrhenian Sea in Figure 9. An exception is the NAd, in proximity of the Po River, where the model simulated fresher waters than observed. Nitrate, chlorophyll, and dissolved oxygen also reached the highest errors in the NAd (Figure 9); this was a direct consequence of using approximate climatological riverine discharges.

Figure 10 quantifies the systematic model overestimation of nitrate in coastal areas and in the NAd in particular. In the NAd, our errors in nutrient estimates are comparable to current ones by operational systems (e.g., Teruzzi et al., 2018), pointing to better riverine representation as a much needed development in Mediterranean models. On the other hand, the reanalysis was skilled in representing the nitrate distribution in open ocean sites, where the simulation and in situ observations had comparable distributions in both surface and deep waters (Figure 10).

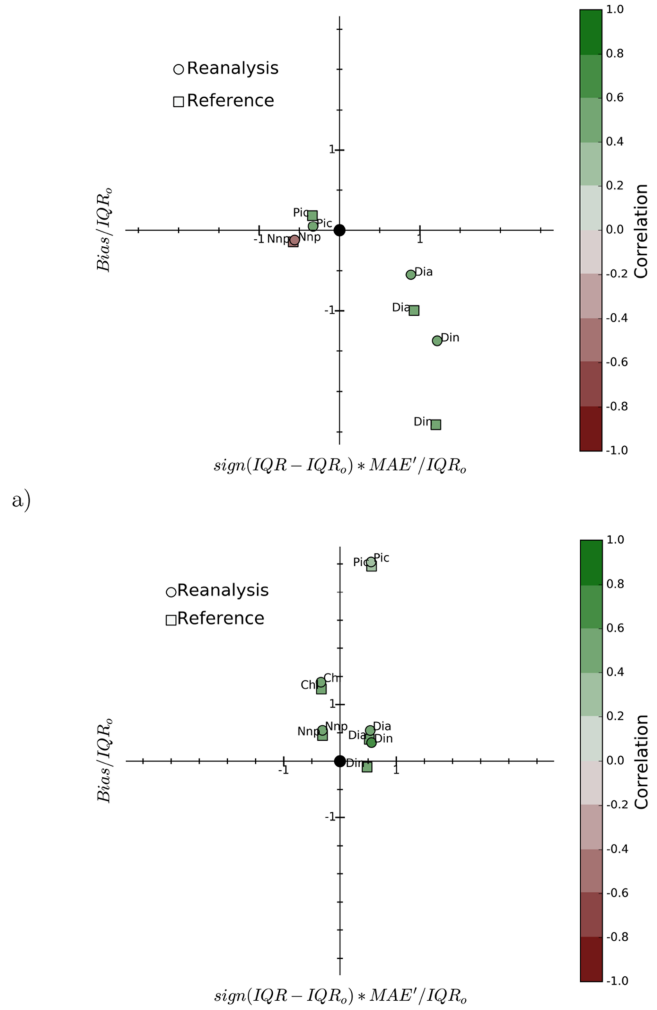

Figure 7. Skill of the reanalysis (circles) in representing the spatial distributions of (a) the ocean color phytoplankton functional type fractions and (b) concentration. The skill of the reference simulation is shown for comparison (squares). IQR = interquartile range of the model estimates; IQRo = interquartile range of the observations; MAE′ = unbiased median absolute error; Bias = median value of the model-to-data deviations. The notation of the variables is explained in Table 2.

The model was skilled in simulating phosphate observations in the NAd (Figures 9 and 10). This mitigated the impact of nitrate overestimation on primary production in the NAd, given that phosphate was the limiting factor in our simulation. The model tended to overestimate the phosphate data in other coastal areas and in the open ocean (Figure 10), but too few data were available for a robust assessment (Figure 9).

Chlorophyll error did not have clear spatial patterns, since high and low error data points coexisted in the same area (e.g., in the Algerian Sea, the Gulf of Lion, and the NAd, Figure 9). Figure 10 shows that the reanalysis slightly overestimated the in situ observations in the open surface water (depths 0–10 m), where the model tended to overestimate also the ocean color product (Figure 3). At the same time, the reanalysis distribution in the deep open ocean (depths 10–200 m) overlapped with the lowest tail of the observed distribution, meaning that the reanalysis tended to underestimate the deep chlorophyll maximum values. In the coastal areas, the simulated and observed distributions are comparable and both reached the highest values in the North Adriatic Sea. Here the model did not produce systematic chlorophyll outliers despite the nitrogen overestimation because simulated phytoplankton production was limited by phosphate (Figure 10). If chlorophyll data from the NAd are removed from our comparison, the reanalysis and observation distributions in coastal areas were 1 order of magnitude lower ("No NAd" in Figure 10).

The simulated PFT concentrations overlapped with the observations, though the reanalysis was skewed toward low values (Tables 2 and S4). However, the relatively low number of sampling stations affects the robustness of the assessment (Figure 9).

3.2.1. Comparison With the Reference Simulation

The reanalysis skill in matching the in situ observations of the biogeochemical variables was not significantly different from the skill of the reference model without assimilation (Figure 8 and Table 2). For the PFT concentrations, assimilation reduced the reference bias versus in situ observations (in particular that of PIC), but it also increased their MAE′ slightly (Figure 8). The reanalysis skill overlapped with that of the reference for nutrients and dissolved oxygen, while phosphate was degraded (Figure 8). The limited impact of PFT assimilation on fitting the in situ data of these biogeochemical variables was expected, because none of them was included in the EnKF analysis (see section 2.2.1), but could only be updated in the forecast model integration (see to this regard the discussion in Pradhan et al., 2019). The relatively low frequency of the assimilative analyses and the relative

Table 1
Skill Metrics for the Reanalysis (Analysis and 1-Month Forecast) and Reference Estimates of the Ocean Color Phytoplankton Functional Type and Total Chlorophyll Concentrations

	R (unitless)			RMSD (mg/m^3)		
Variables	Analysis	Forecast	Reference	Analysis	Forecast	Reference
Diatoms	0.58	0.29	0.26	0.11	0.13	0.13
Dinoflagellates	0.59	0.48	0.41	0.04	0.03	0.03
Nanophytoplankton	0.53	0.32	0.29	0.11	0.13	0.13
Picophytoplankton	0.40	0.27	0.25	0.19	0.20	0.20
Total chlorophyll	0.55	0.32	0.28	0.35	0.41	0.41

Note. The metrics are time averages of the spatial R and RMSD computed at each month of years 1998–2014 in the whole Mediterranean Sea. *R* = correlation; RMSD = root-mean-square deviation.

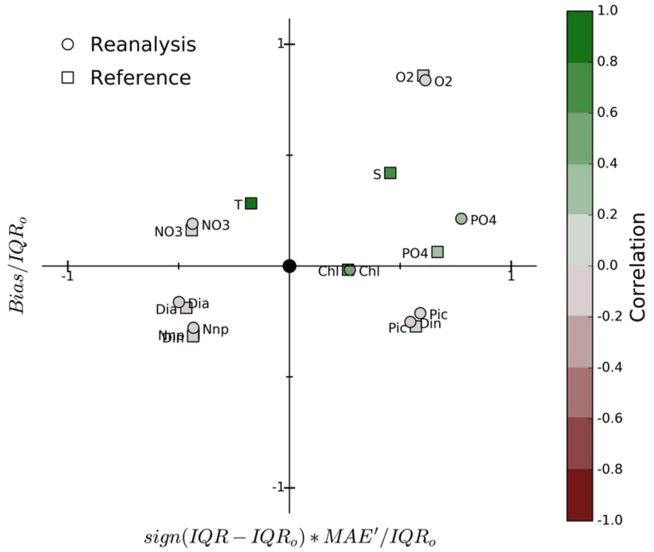

Figure 8. Skill of the reanalysis in simulating in situ observations of physical and biogeochemical variables (circles). The skill of the reference simulation is reported for comparison (squares). IQR = interquartile range of the model estimates; IQR_o = interquartile range of the observations; MAE′ = unbiased median absolute error; Bias = median value of the model-to-data deviations. The notation of the variables is explained in Table 2. Reference and reanalysis skill for T and S were identical because of the one-way coupling of biogeochemistry to physics in the model applied here.

high error of the ocean color PFT data can also help explain this outcome (see Ciavatta et al., 2018). The degradation of phosphate can be related to the analysis increments of phosphorus content of phytoplankton, which acted as a source of nutrients to the system, exacerbating further the reference bias (see section 3.1.1 and Figure S2 in the supporting information). We note that deterioration of nutrients can occur when assimilating ocean color, in particular when issues in the model lead to systematic biases in both the phytoplankton biomass and the nutrient concentrations (see discussion in Gregg et al., 2009), such as overestimation of nutrients inputs and chlorophyll concentration in our case. For example, assimilation of ocean color chlorophyll worsened estimates of nitrate in the NAd in the recent simulations of both Tsiaras et al. (2017) and Teruzzi et al. (2018), in relation to the systematic underestimation of both chlorophyll and nutrients in that region. Though we consider the model suitable to assess the PFT data assimilation strategy in the context of this work (see the relatively good match of model and ocean color PFT fractions in Figures 4–7 and with PFT in situ observations in Figure S4), we highlight that a comprehensive assessment of both nitrate and phosphate fluxes in the model needs to be performed in future studies.

3.3. PFT Ecoregions in the Mediterranean Sea

The good skill of the reanalysis in estimating the PFT fractions inferred from ocean color (Figures 4, 5, and 6) and the improvements with respect to the reference simulation (in particular for MIC ratios, Figure 7a) supported the use of the reanalysis output in Figure 4 to define PFT-based ecoregions in the Mediterranean Sea (Table 3 and Figure 11).

The region of PIC dominance (blue in Figure 11) is the largest in the Mediterranean Sea by far. It covers the offshore waters from the East Mediterranean to the Balearic Islands. It also includes the Central Aegean Sea and the central and southern offshore areas of the Adriatic Sea.

The region of NNP relevance (yellow in Figure 11) covers large bands of the Mediterranean coasts, in particular in the Western Mediterranean, Adriatic Sea, and Levantine.

The region of MIC dominance (in Figure 11) is comparatively small and associated with coastal areas. It covers mainly the NAd and the Gulf of Gabès. However, it spreads also along the Italian Adriatic coast, the Libyan coast, and the Nile Delta.

In general terms, the simulated PFT distribution is largely driven by the gradients of surface dissolved inorganic nutrients available for growth (see the phosphate and nitrate distributions in Figure 9). MIC dominates

Table 2
Comparison of the Statistical Distributions of the Reanalysis (REA), Reference (REF), and Observations (OBS) Matchups in the Whole Mediterranean Sea

Variable (units)	Not	REA median (Q1–Q3)	REF median (Q1–Q3)	OBS median (Q1–Q3)	*N*
Temperature (°C)	T	15.7 (15.0–16.8)	As REA median	14.6 (13.8–17.0)	501,247
Salinity (PSU)	S	38.8 (35.8–39.2)	As REA median	38.3 (37.7–38.6)	223,011
Oxygen (mg/L)	O2	7.9 (6.5–9.2)	7.9 (6.6–9.2)	5.7 (4.7–6.9)	84,537
Nitrate (mmol/m^3)	NO3	5.2 (3.0–7.2)	4.9 (2.8–6.8)	2.9 (0.6–6.3)	3,477
Phosphate (mmol/m^3)	PO4	0.08 (0.02–0.2)	0.07 (0.02–0.2)	0.05 (0.03–0.1)	1,702
Total chl (mg Chl/m^3)	Chl	0.004 ($1{\cdot}10^{-3}$–0.3)	0.003 ($1{\cdot}10^{-3}$–0.3)	0.06 ($7{\cdot}10^{-3}$–0.3)	60,414
Diatoms (mg Chl/m^3)	Dia	0.006 ($1{\cdot}10^{-3}$–0.04)	0.004 ($1{\cdot}10^{-3}$–0.03)	0.02 ($1{\cdot}10^{-2}$–0.06)	822
Dinoflag (mg Chl/m^3)	Din	0.001 ($1{\cdot}10^{-3}$–0.02)	0.001 ($1{\cdot}10^{-3}$–0.01)	0.009 ($5{\cdot}10^{-3}$–0.02)	815
Nanophytopl (mg Chl/m^3)	Nnp	0.03 ($6{\cdot}10^{-3}$–0.07)	0.03 ($6{\cdot}10^{-3}$–0.07)	0.08 ($4{\cdot}10^{-2}$–0.2)	774
Picophytopl (mg Chl/m^3)	Pic	0.04 ($2{\cdot}10^{-2}$–0.10)	0.04 ($2{\cdot}10^{-2}$–0.1)	0.07 ($5{\cdot}10^{-2}$–0.1)	824

Note. "Not" is the notation of the variables used in Figure 8; Q1 and Q3 are the first and third quartiles, respectively; *N* is the number of matchups.

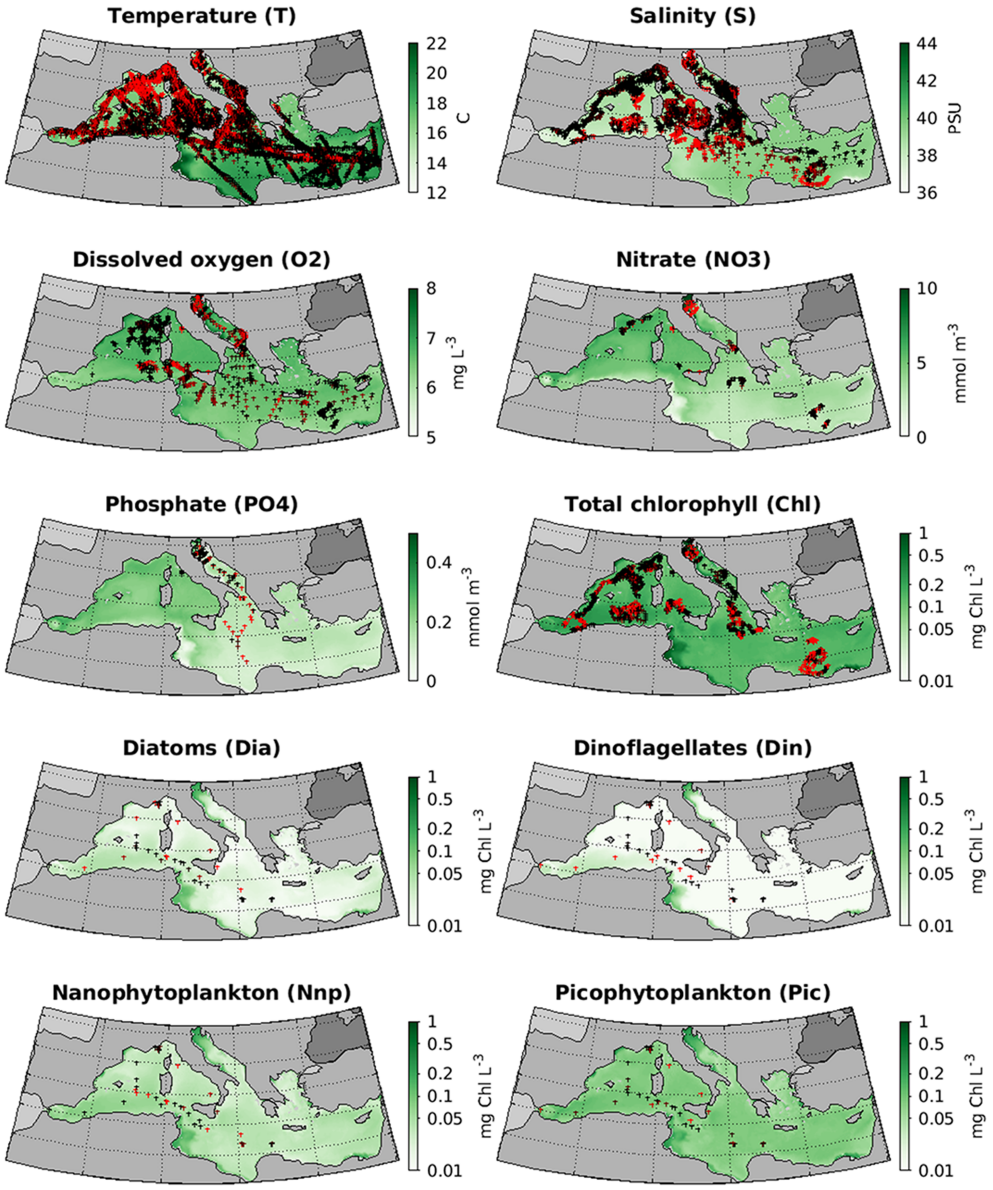

Figure 9. Spatial distribution of the reanalysis errors in simulating the in situ observations collected at Mediterranean sampling sites (crosses). The error is the median absolute difference normalized by the median observation at each site and scaled between the first and third quartiles of the median absolute difference values across sites. The relative errors have increasing values from 0 (low error, black color) to 1 (high error, red color). The maps show the reanalysis average values of the biogeochemical variable in the years 1998–2014.

the areas near the main river mouths, where nutrient abundance can sustain relatively large cells, which require high intracellular nutrient concentration. When nutrient are not limiting, these PFTs tend to accumulate high amount of nutrients (luxury uptake) reducing their availability for smaller phytoplankton. Large phytoplankton is also less heavily grazed with respect to small species: In our model MIC is only preyed on by meso and microzooplankton, while small phytoplankton is preyed on by all the zooplankton functional types, including heterotrophic nanoflagellates. Conversely, the fast-growing, small PIC dominate the large areas of oligotrophic open waters of the Mediterranean. This PFT has a higher surface to volume ratio (due to their smaller size), and therefore, they are able to take up nutrients at lower concentration than large cells. In the model, this feature is described through a parameter representing the affinity for nutrients during uptake (Butenschön et al., 2016); for example, the PIC affinity for phosphorus is double that in diatom. Finally, NNP, which has intermediate characteristics with respect to the other two groups, is relevant

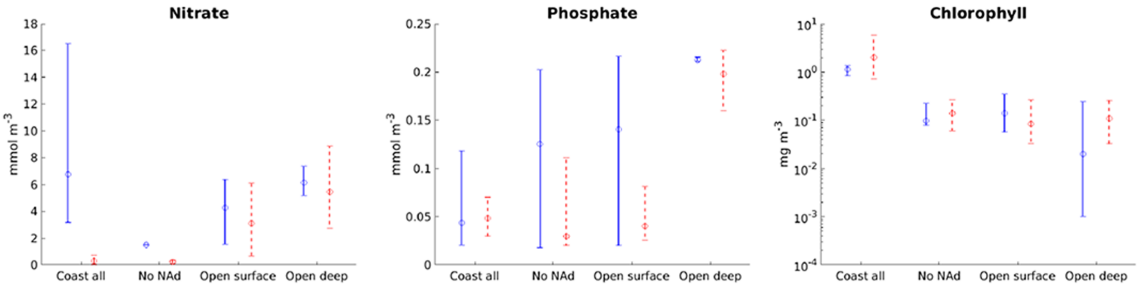

Figure 10. Comparison of the nitrate, phosphate, and chlorophyll reanalysis (blue continuous bars) and in situ data (red dashed bars) in the Mediterranean Sea: coastal, open surface ocean (0–200 m; 0–10 m for chlorophyll) and open deep ocean (200 m to bottom; 10–200 m for chlorophyll). In "No NAd," matchups in the North Adriatic (latitudes higher than 43.5°N) were excluded from the computation of the coastal distribution. The bars represent the first and third quartiles around the median values (circles).

where nutrients are present at intermediate concentrations, as at the edges of the main river plumes and other coastal areas influenced by smaller rivers. Hydrodynamics impacted clearly our simulation and regionalization. For example, the NNP zone expands in the Alborán Sea and beyond because of some oceanic nutrient inputs and mesoscale dynamics associated with the inflow of Atlantic waters, which increases the productivity of those areas (d'Ortenzio & Ribera d'Alcalà, 2009).

The regionalization is based on a reanalysis that matched well the PFT structure inferred from ocean color, but additional considerations might be needed for those areas where the consensus between reanalysis and ocean color was lower. These areas are exemplified by the 75% similarity line in Figure 11, which was reported from the results in Figure 5.

We have confidence in our definition of the MIC region, although the region is included in the coastal areas where the similarity was lower than 75% (Figure 11). In fact, the MIC-dominated areas were overlapping substantially in the reanalysis and ocean color products. The two products differed in the magnitude of the MIC domination, which was noticeably higher from the satellite algorithm than from the model, for example, in the MIC spots in the NAd and in the Gulf of Gabès (Figure 4).

Portions of the PIC region are also bounded by the 75% similarity line, most notably the highly productive Ligurian Sea/Gulf of Lion and the North Aegean Sea (Figure 11). Here NNP was dominant in the ocean color product, while PIC fraction was more than 50% in the reanalysis product (Figure 4). Given the difference in the two products, the consideration of information from in situ PFT observations and local-scale modeling might be needed to confirm the classification of the Ligurian Sea/Gulf of Lion and the North Aegean Sea.

The Ligurian Sea/Gulf of Lion was previously classified as a "blooming" area, based on clustering analysis of the phenology of total chlorophyll from ocean color (d'Ortenzio & Ribera d'Alcalà, 2009; Mayot et al., 2016). The ocean color product applied in our work indicates a climatological NNP blooming or dominance in the region from November to May, and PIC dominance from June to October; on the other hand, our reanalysis simulated NNP dominance only in the month of December (climatological analysis), and PIC dominance in the rest of the year (not shown). The reanalysis and ocean color dissimilarity in representing the PFT seasonal succession and average dominant PFT in the Ligurian Sea/Gulf of Lion is possibly related to our inadequate simulation of the Rhone river inputs (see the highest dissimilarity at the river mouth in Figure 5), also to the

Table 3
Definition of the phytoplankton functional type (PFT) ecoregions in Figure 11

PFT ecoregion	Definition
MIC	Microphytoplankton is dominant: Here microphytoplankton chlorophyll (sum of diatoms and dinoflagellates) ratio to total chlorophyll is higher than the remaining PFT chlorophyll ratios.
NNP	Nanophytoplankton is relevant: Here the nanophytoplankton ratio is the second highest ratio (after picophytoplankton) and summed with micro ratio is higher than the dominant pico ratio.
PIC	Picophytoplankton is absolutely dominant: Here the pico ratio is higher than 50%.

Note. The ecoregion were computed using the surface reanalysis PFT fractions shown in Figure 4.

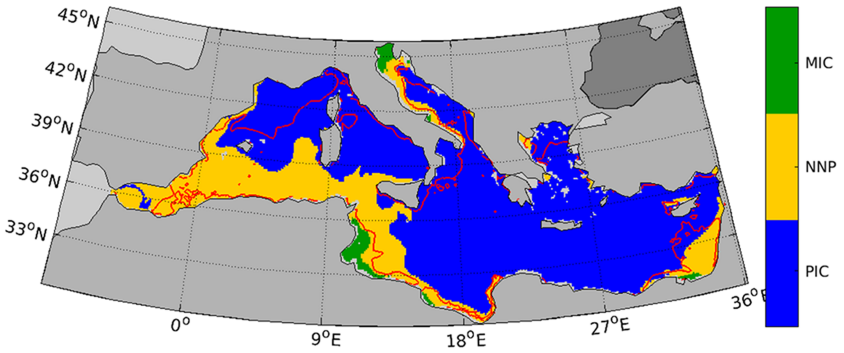

Figure 11. Phytoplankton functional type (PFT)-based ecoregions of the Mediterranean Sea. The PFT ecoregions are defined in Table 3. MIC = microphytoplankton; NNP = nanophytoplankton; PIC = picophytoplankton. The ecoregions were derived from the surface reanalysis PFT fractions shown in Figure 4. The red line represents the similarity value of 75% derived in Figure 5.

misrepresentation of the complex hydrodynamics of the area (see section 3.2). In this region, higher-resolution models might be needed to resolve the hydrodynamic impact on the nutrient transport and zooplankton grazing on the phytoplankton community structure (Auger et al., 2014). However, we note that the satellite model used here was parameterized using high-performance liquid chromatography (HPLC) pigment data (Di Cicco et al., 2017). In the Atlantic Ocean, differences have been seen between size-fractionated chlorophyll estimated through HPLC and through size fractionated filtration (Brewin et al., 2014), with implications for satellite size-fractionated chlorophyll and primary production algorithms (see Brewin et al., 2014; Brewin et al., 2017). For Mediterranean waters, future efforts should be placed on collecting in situ data using other techniques (e.g., flow cytometry, microscopy, and size-fractioned filtration) alongside HPLC, to aid in calibration of future ocean color phytoplankton community structure algorithms.

In the North Aegean Sea, field data suggested that the area is overall dominated by PIC (>60%, Ignatiades et al., 2002; Siokou-Frangou et al., 2002), supporting our reanalysis-based classification (Figure 11). Also, a local-scale ecosystem model confirmed the dominance by PIC in the North Aegean Sea (44%, Tsiaras et al., 2014), while MIC dominance and NNP dominance were driven by nutrient river discharges in confined coastal spots of the North Aegean Sea, for example, the Thermaikos Gulf identified as NNP also in our regionalization (Figure 11).

Finally, the NNP region is bounded by the 75% similarity line in the Alborán Sea, Italian Adriatic coast, and South East Levantine (Figure 11). Here the reanalysis and the ocean color are in general coherent in pointing out NNP in relevant fractions, supporting our classification, though the absolute magnitudes of the fractions can be quite different in the two products (see Figure 4). The different magnitudes are likely related to both the coarse resolution and approximate riverine inputs in our model. Because of that, our model cannot resolve the PFTs as accurately as locally tuned models, which could capture the ecosystem effect of eddies near the Gibraltar Strait (Ramirez-Romero et al., 2014), of the Po input transport (Polimene et al., 2006) and of the Nile plume (e.g., Petihakis et al., 2009). Some in situ observations available in the above areas seem to support their classification in the NNP region. For example, relevant autotrophic NNP fractions were observed in the surface waters of the South East Levantine (comparable to the PIC biomass, Tanaka et al., 2007), in the Alborán Sea (up to 40% of the phytoplankton community, Arin et al., 2002), and in the west coast of the Adriatic Sea (e.g., the largest abundance fractions in Zoppini et al., 1995).

It is worth noting that the MIC and NNP ecoregions in Figure 11 are larger in space than the ecoregions obtainable from the reference simulation (Figure S5 in the supporting information). If compared to ecoregions estimated from the ocean color product, the reanalysis area of NNP relevance is less extended (Figure S6 in the supporting information) and comparable to the area where NNP is dominant in the ocean color product (Figure S7 in the supporting information).

3.4. Regional Plankton-Driven Biogeochemical and Trophic Flows

The magnitude and variability of selected, normalized carbon fluxes, and trophic flows in the Mediterranean PFT ecoregions are presented in Figure 12. Our regionalization approach is supported by these results,

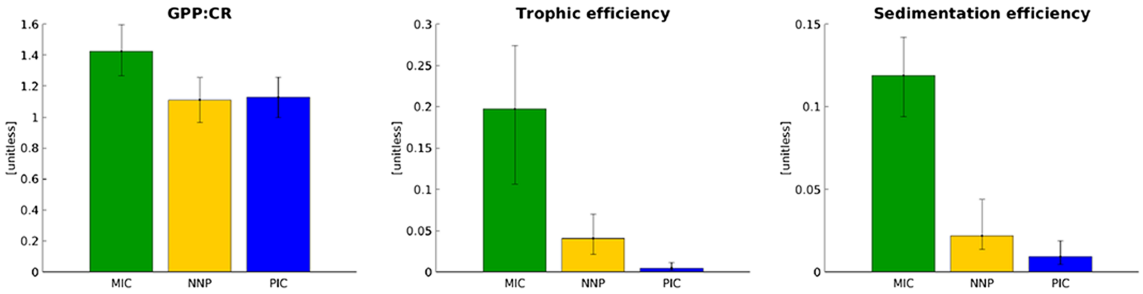

Figure 12. Normalized carbon fluxes in the reanalysis phytoplankton functional type ecoregions in Figure 11 (average values in 1998–2014). The height of the box represents the median value of the region, while the error bar represents the spatial variability within the region (i.e., the interquartile range). Fluxes are defined in section 2.5 and were integrated from the surface to the maximum depth of 200 m. The sedimentation efficiency is an exception, because it was computed at the bottom of the water column. The distributions of the carbon fluxes used in the computation of the efficiencies are shown in Figure S8 in the supporting information. CR = community respiration; GPP = gross primary production; MIC = microphytoplankton; NNP = nanophytoplankton; PIC = picophytoplankton.

because the vertical fluxes were relatively homogeneous within the regions and distinct between regions. In fact, the spatial interquartile ranges were relatively small and not overlapping in Figure 12, with some exceptions.

MIC is the most distinct region, having the maximal and more variable values of all the normalized fluxes. Crucially, the NNP region also diverges from the PIC region in terms of trophic and sedimentation efficiency. That means "relevant" fractions of NNP can make a difference in fluxes in areas dominated by PIC. Apparently, this does not hold for the GPP normalized by the total respiration, since this ratio is not significantly different in NNP and PIC.

The regional differences of the simulated fluxes in Figure 12 can be explained based on our model formulation of the PFT traits. The differences are also determined by "emergent properties" of the ecosystem. These are features that are not coded explicitly in the model but arise from nonlinear synergies of biogeochemical and physical states simulated by the model, such as spatial shifts of the plankton trophic web in relation to nutrient and organic matter availability.

The community production-to-respiration ratio (GPP:CR) is higher than 1 on average, characterizing the epipelagic Mediterranean as an overall autotrophic system (Figure 12). The ratio is higher in the nutrient-rich coastal areas dominated by MIC. In the model, large phytoplankton is preyed by mesozooplankton, supporting the herbivorous food chain and leading to higher transfer efficiency toward upper trophic levels than in regions where only smaller zooplankton are active. This trophic structure implies production of relatively large POC (from fast-sinking large organisms, fecal pellets, grazing, and mortality detritus), which leads to the highest sedimentation efficiency in the MIC region.

At the other extreme, in the oligotrophic PIC region, the median primary production is low and detritus is mainly directed to the dissolved pool (or to small, slowly sinking particles), explaining the smallest sedimentation efficiency. In this region small grazers are advantaged with respect to mesozooplankton, leading to the low efficiency of transfer toward upper trophic levels. These results are coherent with the findings of Allen et al. (2002) and by Lazzari et al. (2016), who estimated that 50% of the potential production is channelled toward the dissolved pool in the Eastern Mediterranean, almost entirely covered by our PIC region.

Normalized fluxes and flows in the NNP region sit between the two other regions (e.g., trophic and sedimentation efficiency), consistent with the intermediate traits of the NNP type (e.g., intermediate parameter values for the mesozooplankton predation and sinking rates in the model).

We argue that the simulated ecoregion efficiencies are meaningful for the real Mediterranean ecosystem because of the substantial correspondence between the reanalysis and ocean color PFT fraction distributions (Figures 4–7), despite some model biases and lack of data to corroborate the carbon flux estimates. In fact, the flows and fluxes in Figure 12 are all normalized quantities; thus, they should be affected only at the second order level by the biases in the phytoplankton biomasses shown in Figure 3. In addition, although flux data were not available to corroborate our novel PFT ecoregions, there is some consistency between flux estimates from our reanalysis and previous works at the scale of the whole Mediterranean Sea. Table 4

Table 4
Average values and uncertainty of selected carbon fluxes in the Mediterranean Sea

Flux (g C·m^{-2}·year^{-1})	Mean	min–max	Sat product[a]	In situ observations[b]	Model[c]
Net production total phyto	69	65.6 ÷ 72.1	68 (65 ÷ 71)	80–90	104
Net production diatoms	7	6.5 ÷ 7.5	12		
Net production dinoflag.	2.3	2.1 ÷ 2.5			
Net production nanophyto	17.6	16.7 ÷ 18.5	33		
Net production picophyto	42	39.8 ÷ 44	23		
Export to sediments	5.2	5 ÷ 5.4			

Note. The uncertainty is presented as minimum–maximum range of the model ensemble in the reanalysis simulation. The fluxes were integrated through the entire water column. Their units are in grams of carbon per square meter per year, and the total flows for the whole Mediterranean Sea can be obtained by multiplying the fluxes by the area of the model domain: $2.4695 \cdot 10^{12}$ m^2. Previous published estimates of the fluxes at the scale of the whole Mediterranean Sea are listed for comparison.
[a]Uitz et al. (2012). [b]Sournia (1973). [c]Di Biagio et al. (2019).

shows that our estimates of the net primary production by the total phytoplankton community matches well the one by Uitz et al. (2012), who used observations from ocean color and in situ campaigns. The confidence bound for our estimates was derived from the range of our 100-member ensemble simulation, and it overlaps with the range provided by Uitz et al. (2012). We note that these estimates are lower than the ones derived from sparse in situ observations (Sournia, 1973) and from other model simulations (Di Biagio et al., 2019).

Our MIC net production is comparable to the one estimated by Uitz et al. (2012) based on their PFT ocean color algorithm, but NNP and PIC production diverge by ~100%. This difference is possibly related to our overestimation of the PIC biomass, in particular in the North West Mediterranean area (Figure 3). However, the PIC production might be underestimated by the ocean color algorithm of Uitz et al. (2012). The Uitz et al. (2012) algorithm is parameterized using pigment data. Differences have been seen in the Atlantic when using ocean color algorithms of size-fractionated primary production tuned using size-fractionated filtration-based methods (Brewin, Tilstone, et al., 2017), with higher estimates of PIC production (see Table 3 of Brewin, Tilstone, et al., 2017). In addition, in situ studies have reported a 65% PIC contribution to the total Mediterranean production (Magazzù & Decembrini, 1995), comparable to 60% we estimated here. Reducing uncertainty in phytoplankton community structure in the Mediterranean Sea will ultimately require more in situ campaigns collecting information on phytoplankton community structure using a range of techniques (e.g., flow cytometry, microscopy, and size-fractioned filtration, alongside HPLC).

Carbon export to the sediment represents 8% of the net primary production in our simulation (Table 4). This value needs to be corroborated in future studies, since other estimates are not available for comparison at the scale of the Mediterranean Sea, to the authors' knowledge.

4. Conclusions

The Mediterranean Sea can be subdivided into three PFT ecoregions, which are homogeneous in the surface phytoplankton community structure as well as in selected, vertically integrated carbon and trophic fluxes:

1. a relatively small region of microphytoplankton dominance, in the most productive coastal areas, where the flows of food and POC toward the upper trophic levels and the sediments, respectively, reach the highest efficiency;
2. a large region of picophytoplankton dominance, covering mainly oligotrophic open ocean waters, which is the less efficient in transferring food and POC;
3. a region where nanophytoplankton is relevant, in coastal and Atlantic-influenced waters, where the efficiency of the carbon fluxes is intermediate.

The regionalization was based on a skilled reanalysis of the phytoplankton community structure of the Mediterranean Sea (1998–2014). The reference simulation of the community structure (i.e., PFT fractions) was improved by the assimilation of a regional, quality-assessed ocean color PFT product, confirming the usefulness of this approach for the reanalysis of marine ecosystems (Ciavatta et al., 2018). However, the reanalysis did not improve significantly the simulation of in situ data of not-assimilated biogeochemical variables, which were not included in the multivariate analysis, and deteriorated the simulation of phosphate, as a consequence of the analysis leading to an increase in the phytoplankton nutrient contents. The

application of mass-conserving assimilative schemes preventing nutrient divergences could be useful to mitigate the issue (e.g., Hemmings et al., 2008), but model improvements to reduce the reference nutrient biases and thus increase the assimilation efficacy are certainly a priority. Such possible improvements include the coupling of higher-resolution physical models or reanalysis (e.g., Adani et al., 2011; Teruzzi et al., 2018), more realistic representation of riverine discharges and improved parameterization of the model PFT chlorophyll-to-carbon ratios.

We acknowledge the limits of our regionalization, which is based on time averaged, surface PFT chlorophyll estimates, and therefore cannot represent the seasonal, interannual, and vertical variability of the plankton community structure. These choices were functional to provide a concise regionalization of the Mediterranean, corroborated by surface satellite data, and characterized by homogenous, vertically integrated ecosystem fluxes.

Our regionalization can add a useful layer of information for basin-scale planning of ecosystem studies, management actions, and sustainable exploitation of the Mediterranean Sea, for example, in the context of the European Marine Strategy Framework Directive (Ayata et al., 2018). For example, one could speculate that the MIC region identified here is the most suitable for fisheries or aquaculture activities sustained by large and efficient particulate organic matter flows (e.g., shellfish production). Then, local ecosystem, high trophic level, and aquaculture models (e.g., Brigolin et al., 2014; Petihakis et al., 2012; Politikos et al., 2015; Tsiaras et al., 2014) could be nested off-line to our reanalysis for fine-scale analysis and aquaculture planning within the sustainable macroregions identified in this work. Such kind of application is the object of our current research in the framework of the European H2020 TAPAS Project.

Acknowledgments
This work has received funding from the European Union's Horizon 2020 research andinnovation programme under grant agreement No 678396, project TAPAS (Tools forAssessment and Planning of Aquaculture Sustainability) This work builds on the Copernicus Marine Environment Monitoring Service (CMEMS) "Optical data modelling and assimilation" (OPTIMA) project, "Toward Operational Size-class Chlorophyll Assimilation (TOSCA) project, and the North-West European Shelf Monitoring and Prediction Service (NOWMAPS) project. CMEMS is implemented by MERCATOR OCEAN in the framework of a delegation agreement with the European Union. This work was also supported by the U.K. NERC through the National Centre for Earth Observation (NCEO), single centre national capability programme – Climate Linked Atlantic Sector Science (NE/R015953/1), the Atlantic BiogeoChemical (ABC) Fluxes Project, and CAMPUS project. The work used the U.K. National Supercomputing Service "ARCHER." Ocean color products were provided by the CMEMS Ocean Colour Thematic Assembly Centre using an ad hoc configuration of the ESA OC-CCI (European Space Agency-Ocean Color Climate Change Initiative, www.esa-oceancolour-cci.org). They were produced by the Global Ocean Satellite monitoring and marine ecosystem studies group (GOS) of the Italian National Research Council (CNR). The regional PFT satellite algorithms were developed in the framework of the Italian Flagship Project RITMARE (la Ricerca ITaliana per il MARE). Part of the in situ observations used here was provided through OGS/NODC infrastructure (http://nodc.ogs.trieste.it). Biogeochemical-ARGO data used here were collected and made freely available by the International Argo Program and the national programs that contribute to it (http://www.argo.ucsd.edu, http://argo.jcommops.org). The Argo Program is part of the Global Ocean Observing System. We also acknowledge the SeaBASS archive for the in situ biooptical data set, freely available at https://seabass.gsfc.nasa.gov. The GHRSST global Level 4 AVHRR_OI product was downloaded from https://podaac.jpl.nasa.gov/dataset/AVHRR_OI-NCEI-L4-GLOB-v2.0. The reanalysis output produced in this work is freely available at the PML Marine System Modelling Group data portal https://portal.ecosystem-modelling.pml.ac.uk. The model ERSEM applied in this work is freely available at https://www.pml.ac.uk/

References

Adani, M., Dobricic, S., & Pinardi, N. (2011). Quality assessment of a 1985–2007 Mediterranean Sea reanalysis. *Journal of Atmospheric and Oceanic Technology*, *28*(4), 569–589. https://doi.org/10.1175/2010JTECHO798.1

Allen, J. I., Blackford, J. C., & Radford, P. J. (1998). An 1-D vertically resolved modelling study of the ecosystem dynamics of the middle and southern Adriatic Sea. *Journal of Marine Systems*, *18*(1-3), 265–286. https://doi.org/10.1016/S0924-7963(98)00015-3

Allen, J. I., Eknes, M., & Evensen, G. (2003). An Ensemble Kalman Filter with a complex marine ecosystem model: Hindcasting phytoplankton in the Cretan Sea. *Annales de Geophysique*, *21*(1), 399–411. https://doi.org/10.5194/angeo-21-399-2003

Allen, J. I., Somerfield, P. J., & Siddorn, J. (2002). Primary and bacterial production in the Mediterranean Sea: A modelling study. *Journal of Marine Systems*, *33*, 473–495.

Arin, L., Morán, X. A. G., & Estrada, M. (2002). Phytoplankton size distribution and growth rates in the Alboran Sea (SW Mediterranean): Short term variability related to mesoscale hydrodynamics. *Journal of Plankton Research*, *24*(10), 1019–1033. https://doi.org/10.1093/plankt/24.10.1019

Auger, P. A., Ulses, C., Estournel, C., Stemmann, L., Somot, S., & Diaz, F. (2014). Interannual control of plankton communities by deep winter mixing and prey/predator interactions in the NW Mediterranean: Results from a 30-year 3D modeling study. *Progress in Oceanography*, *124*, 12–27. https://doi.org/10.1016/j.pocean.2014.04.004

Ayata, S. D., Irisson, J. O., Aubert, A., Berline, L., Dutay, J. C., Mayot, N., et al. (2018). Regionalisation of the Mediterranean basin, a MERMEX synthesis. *Progress in Oceanography*, *163*, 7–20. https://doi.org/10.1016/j.pocean.2017.09.016

Barbieux, M., Organelli, E., Claustre, H., Schmechtig, C., Poteau, A., Boss, E., et al. (2017). A global database of vertical profiles derived from Biogeochemical Argo float measurements for biogeochemical and bio-optical applications, SEANOE, https://doi.org/10.17882/49388

Baretta, J. W., Ebenhöh, W., & Ruardij, P. (1995). The European regional seas ecosystem model, a complex marine ecosystem model. *Netherlands Journal of Sea Research*, *33*(3-4), 233–246. https://doi.org/10.1016/0077-7579(95)90047-0

Baretta-Bekker, J. G., Baretta, J. W., & Ebenhöh, W. (1997). Microbial dynamics in the marine ecosystem model ERSEM II with decoupled carbon assimilation and nutrient uptake. *Journal of Sea Research*, *38*(3-4), 195–211. https://doi.org/10.1016/S1385-1101(97)00052-X

Barlow, R. G., Mantoura, R. F. C., Gough, M. A., & Fileman, T. W. (1993). Pigment signatures of the phytoplankton composition in the northeastern Atlantic during the 1990 spring bloom. *Deep Sea Research, Part II*, *40*(1-2), 459–477. https://doi.org/10.1016/0967-0645(93)90027-K

Basterretxea, G., Font-Muñoz, J. S., Salgado-Hernanz, P. M., Arrieta, J., & Hernández-Carrasco, I. (2018). Patterns of chlorophyll interannual variability in Mediterranean biogeographical regions. *Remote Sensing of Environment*, *215*, 7–17. https://doi.org/10.1016/j.rse.2018.05.027

Blackford, J. C. (1997). An analysis of benthic biological dynamics in a North Sea ecosystem model. *Journal of Sea Research*, *38*(3-4), 213–230. https://doi.org/10.1016/S1385-1101(97)00044-0

Blackford, J. C. (2002). The influence of microphytobenthos on the Northern Adriatic ecosystem: A modelling study. *Estuarine, Coastal and Shelf Science*, *55*(1), 109–123. https://doi.org/10.1006/ecss.2001.0890

Blackford, J. C., & Gilbert, F. J. (2007). pH variability and CO_2 induced acidification in the North Sea. *Journal of Marine Systems*, *64*(1-4), 229–241. https://doi.org/10.1016/j.jmarsys.2006.03.016

Bopp, L., Aumont, O., Cadule, P., Alvain, S., & Gehlen, M. (2005). Response of diatoms distribution to global warming and potential implications: A global model study. *Geophysical Research Letters*, *32*, L19606. https://doi.org/10.1029/2005GL023653

Brewin, R. J., Ciavatta, S., Sathyendranath, S., Jackson, T., Tilstone, G., Curran, K., et al. (2017). Uncertainty in ocean-color estimates of chlorophyll for phytoplankton groups. *Frontiers in Marine Science*, *4*, 104.

Brewin, R. J., Sathyendranath, S., Hirata, T., Lavender, S. J., Barciela, R. M., & Hardman-Mountford, N. J. (2010). A three-component model of phytoplankton size class for the Atlantic Ocean. *Ecological Modelling*, *221*(11), 1472–1483. https://doi.org/10.1016/j.ecolmodel.2010.02.014

Modelling_at_PML/Models/ERSEM. The authors thank Dr. Kostas Tsiaras, Dr. Dionysios Raitsos-Exarchopoulos and Dr. Helen Powley for the useful discussion of the results presented in this paper.

Brewin, R. J. W., Devred, E., Sathyendranath, S., Lavender, S. J., & Hardman-Mountford, N. J. (2011). Model of phytoplankton absorption based on three size classes. *Applied Optics*, *50*, 4353–4364.

Brewin, R. J. W., Sathyendranath, S., Lange, P. K., & Tilstone, G. (2014). Comparison of two methods to derive the size-structure of natural populations of phytoplankton. *Deep-Sea Research Part I*, *85*, 72–79. https://doi.org/10.1016/j.dsr.2013.11.007

Brewin, R. J. W., Sathyendranath, S., Tilstone, G., Lange, P. K., & Platt, T. (2014). A multicomponent model of phytoplankton size structure. *Journal of Geophysical Research: Oceans*, *119*, 3478–3496. https://doi.org/10.1002/2014JC009859

Brewin, R. J. W., Tilstone, G. H., Jackson, T., Cain, T., Miller, P., Lange, P. K., et al. (2017). Modelling size-fractionated primary production in the Atlantic Ocean from remote sensing. *Progress in Oceanography*, *158*, 130–149. https://doi.org/10.1016/j.pocean.2017.02.002

Brigolin, D., Meccia, V. L., Venier, C., Tomassetti, P., Porrello, S., & Pastres, R. (2014). Modelling biogeochemical fluxes across a Mediterranean fish cage farm. *Aquaculture Environment Interactions*, *5*(1), 71–88. https://doi.org/10.3354/aei00093

Burchard, H., Holding, K., & Villareal, M. R. (1999). GOTM, a general ocean turbulence model. Theory, applications and test cases, Tech. Rep. EUR 18745 EN, European Commission.

Butenschön, M., Clark, J., Aldridge, J. N., Allen, J. I., Artioli, Y., Blackford, J., et al. (2016). *Geoscientific Model Development*, *9*(4), 1293–1339. https://doi.org/10.5194/gmd-9-1293-2016

Chisholm, S. W. (1992). Phytoplankton size. In P. G. Falkowski, & A. D. Woodhead (Eds.), *Primary productivity and biogeochemical cycles in the sea*, (pp. 213–237). New York, NY: Plenum Press. https://doi.org/10.1007/978-1-4899-0762-2_12

Ciavatta, S., Brewin, R. J. W., Sk_akala, J., Polimene, L., de Mora, L., Artioli, Y., & Allen, J. I. (2018). Assimilation of ocean-color plankton functional types to improve marine ecosystem simulations. *Journal of Geophysical Research: Oceans*, *123*, 834–854. https://doi.org/10.1002/2017JC013490

Ciavatta, S., Kay, S., Saux-Picart, S., Butenschön, M., & Allen, J. I. (2016). Decadal reanalysis of biogeochemical indicators and fluxes in the North West European shelf-sea ecosystem. *Journal of Geophysical Research: Oceans*, *121*, 1824–1845. https://doi.org/10.1002/2015JC011496

Ciavatta, S., Torres, R., Saux-Picart, S., & Allen, J. I. (2011). Can ocean color assimilation improve biogeochemical hindcasts in shelf seas? *Journal of Geophysical Research*, *116*, C12043. https://doi.org/10.1029/2011JC007219

Cozzi, S., & Giani, M. (2011). River water and nutrient discharges in the Northern Adriatic Sea: Current importance and long term changes. *Continental Shelf Research*, *31*(18), 1881–1893. https://doi.org/10.1016/j.csr.2011.08.010

Cury, P. M., Shin, Y. J., Planque, B., Durant, J. M., Fromentin, J. M., Kramer-Schadt, S., et al. (2008). Ecosystem oceanography for global change in fisheries. *Trends in Ecology & Evolution*, *23*(6), 338–346. https://doi.org/10.1016/j.tree.2008.02.005

Dai, A., Qian, T., Trenberth, K. E., & Milliman, J. D. (2009). Changes in continental freshwater discharge from 1948 to 2004. *Journal of Climate*, *22*(10), 2773–2792. https://doi.org/10.1175/2008JCLI2592.1

D'Alimonte, D., Mélin, F., Zibordi, G., & Berthon, J.-F. (2003). Use of the novelty detection technique to identify the range of applicability of the empirical ocean color algorithms. *IEEE Transactions on Geoscience and Remote Sensing*, *41*(12), 2833–2843. https://doi.org/10.1109/TGRS.2003.818020

D'Alimonte, D., & Zibordi, G. (2003). Phytoplankton determination in an optically complex coastal region using a multilayer perceptron neural network. *IEEE Transactions on Geoscience and Remote Sensing*, *41*(12), 2861–2868. https://doi.org/10.1109/TGRS.2003.817682

de Fommervault, O. P., D'Ortenzio, F., Mangin, A., Serra, R., Migon, C., Claustre, H., et al. (2015). Seasonal variability of nutrient concentrations in the Mediterranean Sea: Contribution of Bio-Argo floats. *Journal of Geophysical Research: Oceans*, *120*, 8528–8550. https://doi.org/10.1002/2015JC011103

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597. https://doi.org/10.1002/qj.828

Devred, E., Sathyendranath, S., Stuart, V., Maas, H., Ulloa, O., & Platt, T. (2006). A two-component model of phytoplankton absorption in the open ocean: Theory and applications. *Journal of Geophysical Research*, *111*, C03011. https://doi.org/10.1029/2005jc002880

Di Biagio, V., Cossarini, G., Salon, S., Lazzari, P., Querin, S., Sannino, G., & Solidoro, C. (2019). Temporal scales of variability in the Mediterranean Sea ecosystem: Insight from a coupled model. *Journal of Marine Systems*, *197*, 103176. https://doi.org/10.1016/j.jmarsys.2019.05.002

Di Cicco, A., Sammartino, M., Marullo, S., & Santoleri, R. (2017). Regional empirical algorithms for an improved identification of phytoplankton functional types and size classes in the Mediterranean Sea using satellite data. *Frontiers in Marine Science*, *4*, 126. https://doi.org/10.3389/fmars.2017.00126

d'Ortenzio, F., & Ribera d'Alcalà, M. (2009). On the trophic regimes of the Mediterranean Sea: A satellite analysis. *Biogeosciences*, *6*(2), 139–148. https://doi.org/10.5194/bg-6-139-2009

Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte-Carlo methods to forecast error statistics. *Journal of Geophysical Research*, *99*(C5), 10143–10162. https://doi.org/10.1029/94JC00572

Evensen, G. (2003). The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dynamics*, *53*(4), 343–367. https://doi.org/10.1007/s10236-003-0036-9

Ford, D., & Barciela, R. (2017). Global marine biogeochemical reanalyses assimilating two different sets of merged ocean colour products. *Remote Sensing of Environment*, *203*, 40–54. https://doi.org/10.1016/j.rse.2017.03.040

Geider, R. J., MacIntyre, H. L., & Kana, T. M. (1997). Dynamic model of phytoplankton growth and acclimation: Responses of the balanced growth rate and the chlorophyll a:carbon ratio to light, nutrient-limitation and temperature. *Marine Ecology Progress Series*, *148*(1-3), 187–200. https://doi.org/10.3354/meps148187

Gieskes, W. W. C., Kraay, G. W., Nontji, A., & Setiapermana, D. (1988). Monsoonal alternation of a mixed and a layered structure in the phytoplankton of the euphotic zone of the Banda Sea (Indonesia): A mathematical analysis of algal pigment fingerprints. *Netherlands Journal of Sea Research*, *22*(2), 123–137. https://doi.org/10.1016/0077-7579(88)90016-6

Gregg, W. W., Friedrichs, M. A., Robinson, A. R., Rose, K. A., Schlitzer, R., Thompson, K. R., & Doney, S. C. (2009). Skill assessment in ocean biological data assimilation. *Journal of Marine Systems*, *76*(1-2), 16–33. https://doi.org/10.1016/j.jmarsys.2008.05.006

Hallegraeff, G. M. (2003). Harmful algal blooms: A global overview. *Manual on Harmful Marine Microalgae*, *33*, 1–22.

Hansell, D. A. (2013). Recalcitrant dissolved organic carbon fractions. *Marine Sciences*, *5*(1), 421–445. https://doi.org/10.1146/annurev-marine-120710-100757

Hemmings, J. C., Barciela, R. M., & Bell, M. J. (2008). Ocean color data assimilation with material conservation for improving model estimates of air-sea CO2 flux. *Journal of Marine Research*, *66*(1), 87–126. https://doi.org/10.1357/002224008784815739

Hirata, T., Hardman-Mountford, N. J., Brewin, R. J. W., Aiken, J., Barlow, R., Suzuki, K., et al. (2011). Synoptic relationships between surface Chlorophyll-a and diagnostic pigments specific to phytoplankton functional types. *Biogeosciences*, *8*(2), 311–327. https://doi.org/10.5194/bg-8-311-2011

Holt, J., Harle, J., Proctor, R., Michel, S., Ashworth, M., Batstone, C., et al. (2008). Modelling the global coastal ocean. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *367*(1890), 939–951.

Holt, J. T., & James, I. D. (2001). An s coordinate density evolving model of the northwest European continental shelf: 1. Model description and density structure. *Journal of Geophysical Research*, *106*(C7), 14015–14034. https://doi.org/10.1029/2000JC000304

Holt, J. T., Proctor, R., Blackford, J. C., Allen, J. I., & Ashworth, M. (2004). Advective controls on primary production in the stratified western Irish Sea: An eddy-resolving model study. *Journal of Geophysical Research*, *109*, C05024. https://doi.org/10.1029/2003jc001951

Hu, J., Fennel, K., Mattern, J. P., & Wilkin, J. (2012). Data assimilation with a local ensemble Kalman filter applied to a three-dimensional biological model of the Middle Atlantic Bight. *Journal of Marine Systems*, *94*, 145–156. https://doi.org/10.1016/j.jmarsys.2011.11.016

Ignatiades, L., Psarra, S., Zervakis, V., Pagou, K., Souvermezoglou, E., Assimakopoulou, G., & Gotsis-Skretas, O. (2002). Phytoplankton size-based dynamics in the Aegean Sea (Eastern Mediterranean). *Journal of Marine Systems*, *36*(1-2), 11–28. https://doi.org/10.1016/S0924-7963(02)00132-X

IOCCG (2014). Phytoplankton functional types from space. In (Reports of the International Ocean-Colour Coordinating Group (IOCCG); 15) Sathyendranath, S., Aiken, J., Alvain, S., Barlow, R., Bouman, H., Bracher, A., Brewin, R., Bricaud, A., Brown, C.W., Ciotti, A.M. and Clementson, L.A., 2014. (pp. 1-156). International Ocean-Colour Coordinating Group.

James, I. D. (1996). Advection schemes for shelf sea models. *Journal of Marine Systems*, *8*(3-4), 237–254. https://doi.org/10.1016/0924-7963(96)00008-5

Janjić, T., McLaughlin, D., Cohn, S. E., & Verlaan, M. (2014). Conservation of mass and preservation of positivity with ensemble-type Kalman filter algorithms. *Monthly Weather Review*, *142*(2), 755–773. https://doi.org/10.1175/MWR-D-13-00056.1

Kalaroni, S., Tsiaras, K., Petihakis, G., Hoteit, I., Economou-Amilli, A., & Triantafyllou, G. (2016). Data assimilation of depth-distributed satellite chlorophyll-α in two Mediterranean contrasting sites. *Journal of Marine Systems*, *160*, 40–53. https://doi.org/10.1016/j.jmarsys.2016.03.018

Kay, S., Andersson, H., Catalan, I., Eilola, K., Jordà, G., Ramirez-Romero, E., Wehde, W. (2018) Projections of physical and biogeochemical parameters and habitat indicators for European seas, including synthesis of Sea Level Rise and storminess. H2020 CERES project deliverable D1.3, https://ceresproject.eu/

Kay, S., & Butenschön, M. (2018). Projections of change in key ecosystem indicators for planning and management of marine protected areas: An example study for European seas. *Estuarine, Coastal and Shelf Science*, *201*, 172–184. https://doi.org/10.1016/j.ecss.2016.03.003

Kostadinov, T. S., Cabré, A., Vedantham, H., Marinov, I., Bracher, A., Brewin, R. J. W., et al. (2017). Intercomparison of phytoplankton functional types derived from ocean color algorithms and earth system models: Phenology. *Remote Sensing of Environment*, *190*, 162–177. https://doi.org/10.1016/j.rse.2016.11.014

Krom, M. D., Kress, N., Brenner, S., & Gordon, L. I. (1991). Phosphorus limitation of primary productivity in the eastern Mediterranean Sea. *Limnology and Oceanography*, *36*(3), 424–432. https://doi.org/10.4319/lo.1991.36.3.0424

Kwiatkowski, L., Aumont, O., & Bopp, L. (2019). Consistent trophic amplification of marine biomass declines under climate change. *Global Change Biology*, *25*(1), 218–229. https://doi.org/10.1111/gcb.14468

Lavigne, H., D'ortenzio, F., d'Alcalà, M. R., Claustre, H., Sauzède, R., & Gacic, M. (2015). On the vertical distribution of the chlorophyll a concentration in the Mediterranean Sea: A basin-scale and seasonal approach. *Biogeosciences*, *12*(16), 5021–5039. https://doi.org/10.5194/bg-12-5021-2015

Lazzari, P., Solidoro, C., Ibello, V., Salon, S., Teruzzi, A., Béranger, K., et al. (2012). Seasonal and inter-annual variability of plankton chlorophyll and primary production in the Mediterranean Sea: A modelling approach. *Biogeosciences*, *9*(1), 217–233. https://doi.org/10.5194/bg-9-217-2012

Lazzari, P., Solidoro, C., Salon, S., & Bolzon, G. (2016). Spatial variability of phosphate and nitrate in the Mediterranean Sea: A modeling approach. *Deep Sea Research Part I: Oceanographic Research Papers*, *108*, 39–52. https://doi.org/10.1016/j.dsr.2015.12.006

Lee, Z., Marra, J., Perry, M. J., & Kahru, M. (2015). Estimating oceanic primary productivity from ocean color remote sensing: A strategic assessment. *Journal of Marine Systems*, *149*, 50–59. https://doi.org/10.1016/j.jmarsys.2014.11.015

Magazzù, G., & Decembrini, F. (1995). Primary production, biomass and abundance of phototrophic picoplankton in the Mediterranean Sea: A review. *Aquatic Microbial Ecology*, *9*(1), 97–104. https://doi.org/10.3354/ame009097

Mayorga, E., Seitzinger, S. P., Harrison, J. A., Dumont, E., Beusen, A. H. W., Bouwman, A. F., et al. (2010). Global Nutrient Export from WaterSheds 2 (NEWS 2): Model development and implementation. *Environmental Modelling and Software*, *25*(7), 837–853. https://doi.org/10.1016/j.envsoft.2010.01.007

Mayot, N., D'Ortenzio, F., d'Alcala, M. R., Lavigne, H., & Claustre, H. (2016). Interannual variability of the Mediterranean trophic regimes from ocean color satellites. *Biogeosciences*, *13*(6), 1901–1917. https://doi.org/10.5194/bg-13-1901-2016

Natvik, L. J., & Evensen, G. (2003). Assimilation of ocean color data into a biochemical model of the North Atlantic—Part 1. Data assimilation experiments. *Journal of Marine Systems*, *40-41*, 127–153. https://doi.org/10.1016/s0924-7963(03)00016-2

Navarro, G., Almaraz, P., Caballero, I., Vázquez, Á., & Huertas, I. E. (2017). Reproduction of spatio-temporal patterns of major Mediterranean phytoplankton groups from remote sensing OC-CCI data. *Frontiers in Marine Science*, *4*, 246. https://doi.org/10.3389/fmars.2017.00246

Navarro, G., Alvain, S., Vantrepotte, V., & Huertas, I. E. (2014). Identification of dominant phytoplankton functional types in the Mediterranean Sea based on a regionalized remote sensing approach. *Remote Sensing of Environment*, *152*, 557–575. https://doi.org/10.1016/j.rse.2014.06.029

Nerger, L., & Gregg, W. W. (2007). Assimilation of SeaWiFS data into a global ocean-biogeochemical model using a local SEIK filter. *Journal of Marine Systems*, *68*(1-2), 237–254. https://doi.org/10.1016/j.jmarsys.2006.11.009

Petihakis, G., Smith, C. J., Triantafyllou, G., Sourlantzis, G., Papadopoulou, K. N., Pollani, A., & Korres, G. (2007). Scenario testing of fisheries management strategies using a high resolution ERSEM–POM ecosystem model. *ICES Journal of Marine Science*, *64*(9), 1627–1640. https://doi.org/10.1093/icesjms/fsm161

Petihakis, G., Triantafyllou, G., Allen, I. J., Hoteit, I., & Dounas, C. (2002). Modelling the spatial and temporal variability of the Cretan Sea ecosystem. *Journal of Marine Systems*, *36*(3-4), 173–196. https://doi.org/10.1016/S0924-7963(02)00186-0

Petihakis, G., Triantafyllou, G., Koutsoubas, D., Allen, I., & Dounas, C. (1999). Modelling the annual cycles of nutrients and phytoplankton in a Mediterranean lagoon (Gialova, Greece). *Marine Environmental Research*, *48*(1), 37–58. https://doi.org/10.1016/S0141-1136(99)00031-8

Petihakis, G., Triantafyllou, G., Tsiaras, K., Korres, G., Pollani, A., & Hoteit, I. (2009). Eastern Mediterranean biogeochemical flux model—Simulations of the pelagic ecosystem. *Ocean Science*, *5*(1), 29–46. https://doi.org/10.5194/os-5-29-2009

Petihakis, G., Tsiaras, K., Triantafyllou, G., Korres, G., Tsagaraki, T. M., Tsapakis, M., et al. (2012). Application of a complex ecosystem model to evaluate effects of finfish culture in Pagasitikos Gulf, Greece. *Journal of Marine Systems*, *94*, S65–S77. https://doi.org/10.1016/j.jmarsys.2011.11.002

Pinardi, N., Allen, I., Demirov, E., De Mey, P., Korres, G., Lascaratos, A., et al. (2003). The Mediterranean ocean forecasting system: First phase of implementation (1998–2001). *Annales Geophysicae*, *21*(1), 3–20. https://doi.org/10.5194/angeo-21-3-2003

Polimene, L., Allen, J. I., & Zavatarelli, M. (2006). Model of interactions between dissolved organic carbon and bacteria in marine systems. *Aquatic Microbial Ecology*, *43*(2), 127–138. https://doi.org/10.3354/ame043127

Polimene, L., Pinardi, N., Zavatarelli, M., & Colella, S. (2006). The Adriatic Sea ecosystem seasonal cycle: Validation of a three-dimensional numerical model. *Journal of Geophysical Research*, *111*, C03S19. https://doi.org/10.1029/2005JC003260

Politikos, D., Somarakis, S., Tsiaras, K. P., Giannoulaki, M., Petihakis, G., Machias, A., & Triantafyllou, G. (2015). Simulating anchovy's full life cycle in the northern Aegean Sea (eastern Mediterranean): A coupled hydro-biogeochemical–IBM model. *Progress in Oceanography*, *138*, 399–416. https://doi.org/10.1016/j.pocean.2014.09.002

Pradhan, H. K., Völker, C., Losa, S. N., Bracher, A., & Nerger, L. (2019). Assimilation of global total chlorophyll OC-CCI data and its impact on individual phytoplankton fields. *Journal of Geophysical Research: Oceans*, *124*, 470–490. https://doi.org/10.1029/2018JC014329

Pujo-Pay, M., Conan, P., Oriol, L., Cornet-Barthaux, V., Falco, C., Ghiglione, J.-F., et al. (2011). Integrated survey of elemental stoichiometry (C, N, P) from the western to eastern Mediterranean Sea. *Biogeosciences*, *8*(4), 883–899. https://doi.org/10.5194/bg-8-883-2011

Ramirez-Romero, E., Vichi, M., Castro, M., Macías, J., Macías, D., García, C. M., & Bruno, M. (2014). Modeling the biogeochemical seasonal cycle in the Strait of Gibraltar. *Journal of Marine Systems*, *139*, 348–361. https://doi.org/10.1016/j.jmarsys.2014.07.017

Reynolds, R. W., Smith, T. M., Liu, C., Chelton, D. B., Casey, K. S., & Schlax, M. G. (2007). Daily high-resolution blended analyses for sea surface temperature. *Journal of Climate*, *20*(22), 5473–5496. https://doi.org/10.1175/2007JCLI1824.1

Ribera d'Alcalà, M., Civitarese, G., Conversano, F., & Lavezza, R. (2003). Nutrient ratios and fluxes hint at overlooked processes in the Mediterranean Sea. *Journal of Geophysical Research*, *108*(C9), 8106. https://doi.org/10.1029/2002jc001650

Sammartino, M., Di Cicco, A., Marullo, S., & Santoleri, R. (2015). Spatiotemporal variability of micro-, nano- and pico-phytoplankton in the Mediterranean Sea from satellite ocean colour data of SeaWiFS. *Ocean Science*, *11*(5), 759–778. https://doi.org/10.5194/os-11-759-2015

Shimoda, Y., & Arhonditsis, G. B. (2016). Phytoplankton functional type modelling: Running before we can walk? A critical evaluation of the current state of knowledge. *Ecological Modelling*, *320*, 29–43. https://doi.org/10.1016/j.ecolmodel.2015.08.029

Siokou-Frangou, I., Bianchi, M., Christaki, U., Christou, E. D., Giannakourou, A., Gotsis, O., et al. (2002). Carbon flow in the planktonic food web along a gradient of oligotrophy in the Aegean Sea (Mediterranean Sea). *Journal of Marine Systems*, *33*, 335–353.

Skákala, J., Ford, D., Brewin, R. J., McEwan, R., Kay, S., Taylor, B., et al. (2018). The assimilation of phytoplankton functional types for operational forecasting in the northwest European shelf. *Journal of Geophysical Research: Oceans*, *123*, 5230–5247. https://doi.org/10.1029/2018JC014153

Sournia, A. (1973). La production primaire planctonique en Méditerranée; essai de mise à jour. Cooperative Investigations in the Mediterranean, International Coordinator and Operational Unit; Étude en commun de la Méditerranée, Coordonnateur international et Unité opérationnelle.

Storto, A., Masina, S., & Dobricic, S. (2013). Ensemble spread-based assessment of observation impact: Application to a global ocean analysis system. *Quarterly Journal of the Royal Meteorological Society*, *139*(676), 1842–1862. https://doi.org/10.1002/qj.2071

Tanaka, T., Zohary, T., Krom, M. D., Law, C. S., Pitta, P., Psarra, S., et al. (2007). Microbial community structure and function in the Levantine Basin of the eastern Mediterranean. *Deep Sea Research Part I: Oceanographic Research Papers*, *54*(10), 1721–1743. https://doi.org/10.1016/j.dsr.2007.06.008

Teruzzi, A., Bolzon, G., Salon, S., Lazzari, P., Solidoro, C., & Cossarini, G. (2018). Assimilation of coastal and open sea biogeochemical data to improve phytoplankton simulation in the Mediterranean Sea. *Ocean Modelling*, *132*, 46–60. https://doi.org/10.1016/j.ocemod.2018.09.007

Teruzzi, A., Dobricic, S., Solidoro, C., & Cossarini, G. (2014). A 3-D variational assimilation scheme in coupled transport-biogeochemical models: Forecast of Mediterranean biogeochemical properties. *Journal of Geophysical Research: Oceans*, *119*, 200–217. https://doi.org/10.1002/2013JC009277

Triantafyllou, G., Hoteit, I., & Petihakis, G. (2003). A singular evolutive interpolated Kalman filter for efficient data assimilation in a 3-D complex physical–biogeochemical model of the Cretan Sea. *Journal of Marine Systems*, *40*, 213–231.

Triantafyllou, G., Petihakis, G., Dounas, C., Koutsoubas, D., Arvanitidis, C., & Eleftheriou, A. (2000). Temporal variations in benthic communities and their response to physicochemical forcing: A numerical approach. *ICES Journal of Marine Science*, *57*(5), 1507–1516. https://doi.org/10.1006/jmsc.2000.0923

Tsiaras, K. P., Hoteit, I., Kalaroni, S., Petihakis, G., & Triantafyllou, G. (2017). A hybrid ensemble-OI Kalman filter for efficient data assimilation into a 3-D biogeochemical model of the Mediterranean. *Ocean Dynamics*, *67*(6), 673–690. https://doi.org/10.1007/s10236-017-1050-7

Tsiaras, K. P., Petihakis, G., Kourafalou, V. H., & Triantafyllou, G. (2014). Impact of the river nutrient load variability on the North Aegean ecosystem functioning over the last decades. *Journal of Sea Research*, *86*, 97–109. https://doi.org/10.1016/j.seares.2013.11.007

Uitz, J., Claustre, H., Morel, A., & Hooker, S. B. (2006). Vertical distribution of phytoplanton communities in open ocean: An assessment based on surface chlorophyll. *Journal of Geophysical Research*, *111*, C08005. https://doi.org/10.1029/2005JC003207

Uitz, J., Stramski, D., Gentili, B., D'Ortenzio, F., & Claustre, H. (2012). Estimates of phytoplankton class-specific and total primary production in the Mediterranean Sea from satellite ocean color observations. *Global Biogeochemical Cycles*, *26*, GB2024. https://doi.org/10.1029/2011GB004055

Vichi, M., Zavatarelli, M., & Pinardi, N. (1998). Seasonal modulation of microbially mediated carbon fluxes in the northern Adriatic Sea—A model study. *Fisheries Oceanography*, *7*(3-4), 182–190. https://doi.org/10.1046/j.1365-2419.1998.00082.x

Vidussi, F., Claustre, H., Manca, B. B., Luchetta, A., & Marty, J. C. (2001). Phytoplankton pigment distribution in relation to upper thermocline circulation in the eastern Mediterranean Sea during winter. *Journal of Geophysical Research*, *106*(C9), 19,939–19,956.

Volpe, G., Colella, S., Forneris, V., Tronconi, C., & Santoleri, R. (2012). The Mediterranean ocean colour observing system—System development and product validation. *Ocean Science*, *8*(5), 869–883. https://doi.org/10.5194/os-8-869-2012

Volpe, G., Colella, S., Brando, V. E., Forneris, V., LaPadula, F., Di Cicco, A., Sammartino, M., Bracaglia, M., Artuso, F., & Santoleri, R. (2019). Mediterranean ocean colour Level 3 operational multi-sensor processing. *Ocean Science*, *15*(1), 127–146.

Volpe, G., Santoleri, R., Vellucci, V., Ribera d'Alcala, M., Marullo, S., & D'Ortenzio, F. (2007). The colour of the Mediterranean Sea: Global versus regional bio-optical algorithms evaluation and implication for satellite chlorophyll estimates. *Remote Sensing of Environment*, *107*(4), 625–638. https://doi.org/10.1016/j.rse.2006.10.017

Werdell, P. J., & Bailey, S. W. (2005). An improved in-situ bio-optical data set for ocean color algorithm development and satellite data product validation. *Remote Sensing of Environment*, *98*(1), 122–140. https://doi.org/10.1016/j.rse.2005.07.001

Zavatarelli, M., Baretta, J. W., Baretta-Bekker, J. G., & Pinardi, N. (2000). The dynamics of the Adriatic Sea ecosystem.: An idealized model study. *Deep Sea Research Part I: Oceanographic Research Papers*, *47*(5), 937–970. https://doi.org/10.1016/S0967-0637(99)00086-2

Zoppini, A., Pettine, M., Totti, C., Puddu, A., Artegiani, A., & Pagnotta, R. (1995). Nutrients, standing crop and primary production in western coastal waters of the Adriatic Sea. *Estuarine, Coastal and Shelf Science*, *41*(5), 493–513. https://doi.org/10.1016/0272-7714(95)90024-1